

# nature

THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

## COMPLEX DISEASES GET COMPLICATED

A diet-microbiota-brain- $\beta$ -cell interaction drives  
obesity and the metabolic syndrome **PAGES 185 & 213**

### MISCONDUCT

#### LEARNING THE HARD WAY

Carelessness can land good  
researchers in bad trouble

**PAGE 173**

### MICROBIOLOGY

#### TALKING MICROBIOMES

Chemical cross-talk of  
microbial communities

**PAGE 170**

### PHILOSOPHY

#### 'STEMNESS' DEFINED

Semantic spring-clean offers  
field a way forward

**PAGE 176**

**NATURE.COM/NATURE**

9 June 2016 £10

Vol 534, No. 7606





# THIS WEEK

## EDITORIALS

**RENEWABLES** The future of energy remains clean and green **p.152**

**WORLD VIEW** Use gene-drive concerns to push open science **p.153**



**CURTAIN CALL** Older monkeys struggle with vocal prowess **p.155**

## Humanity's forgotten family

*Hominin fossils discovered near the site of the 'hobbit' *Homo floresiensis* provide yet more evidence that the human lineage is more diverse than was ever imagined.*

Arthur C. Clarke wrote in 2001: *A Space Odyssey* that behind every person now living stand 30 ghosts, for that is by how many the dead outnumber the living. That was in 1968 — the number reckoned today would probably be greater. The human lineage diverged from that of chimpanzees some 5 million to 7 million years ago. Were we able to mark the remains of all our ancestors from that point, the world would be one enormous cemetery.

The most likely fate of any living organism is dissociation into its component molecules, if not reabsorption as food into something else. That makes the chance ineffably remote that the remains of any one individual will be fossilized in any recognizable form, and, this having been achieved, be recognized as such by a passing palaeontologist before the fossil, too, crumbles to dust.

It is possible that many human species once existed, but became extinct with such finality that even those few that were fossilized have since disappeared, leaving absolutely no trace that generations of a distinct species lived and died on this planet — a kind of double extinction, without hope of memorial or discovery. Fossils from the human lineage are scarce, and, given the numbers that must once have lived, the percentage recovered must hardly be significantly different from zero. (You can read about some of those that have been found in our *Nature* collection at [go.nature.com/1zjssjs](http://go.nature.com/1zjssjs).)

### LONG-LOST RELATIONS

Hence the surprise when, in 2004, a group of scientists in Indonesia and Australia announced the discovery of what became known as *Homo floresiensis*, a species of unusual, dwarfed hominin — that is, a creature living or extinct more closely related to us than to chimps — whose remains were found in Liang Bua cave on the island of Flores in Indonesia (P. Brown *et al. Nature* **431**, 1055–1061; 2004).

There was some doubt at the time that *H. floresiensis* represented a real species rather than a variant of modern humans affected by some disease or pathological condition, but this dissent was gradually eroded, not only by a long palaeontological record at Liang Bua, but also by a rich archaeological record in the island's Soa Basin, some distance to the east, showing that hominins of some sort had lived in the region for up to one million years (A. Brumm *et al. Nature* **464**, 748–752; 2010.) Yet direct evidence, in the form of bones and teeth, was elusive. Until now.

On page 245, researchers report a fragment of mandible and six isolated teeth of hominins from Mata Menge in the Soa Basin that they describe as similar to those of *H. floresiensis*, but more primitive in some respects and — if anything — even smaller. In an accompanying paper on page 249, they show that the remains were deposited 700,000 years ago, many thousands of years before those from Liang Bua.

The researchers take the appropriately cautious and parsimonious view that these hominins were most closely related to early Asian *Homo erectus*, on the grounds that this is the only species of hominin otherwise known to have inhabited that part of the world at that time (see

page 164). However, it remains possible, as an accompanying News & Views on page 188 explains, that these creatures might represent some very early, pre-*H. erectus* exodus from Africa. If so, that expands our ignorance from a barely manageable ocean into a gulf of interstellar magnitude, implying that a wholly unknown plethora of hominins lived in Eurasia millions of years earlier than anyone suspected, just one of whose number has been found in the region's southeastern extremity to betray the possibility that such an array of species ever existed.

**“Researchers are less eager than they once were to string fossils together into confident chains of ancestry and descent.”**

Is this unwarranted speculation? Perhaps not: the discovery of *H. floresiensis* prompted a sea change in palaeoanthropologists' attitudes to the unknown. Researchers are less eager than they once were to string fossils together into confident chains of ancestry and descent. They are more likely to reappraise the various oddities of human evolution, no longer dismissing them as fossils that are hard to fit into the current paradigm of ancestry and descent, but seeing them as representatives of entirely unsuspected branches of the human family tree. One thinks of the many hominin remains recovered over the past few decades from China, some of which do not quite fit into any current species. Or of *Homo heidelbergensis*, an increasingly unwieldy catch-all for hominins from the Middle Pleistocene epoch (781,000–126,000 years ago); or of *H. erectus* itself, a grouping of such variety that some have found it hard to accept that all the fossils ascribed to it comprise a single species. And there are others less familiar, such as skulls from Iwo Eleru in Nigeria that look surprisingly archaic, despite being assigned a relatively recent date of as few as 11,700 years ago (K. Harvati *et al. PLoS ONE* **6**, e24024; 2011).

Studies on human DNA, ancient and modern, have reinforced this trend. The recovery of an entire genome of a hitherto unknown archaic hominin from a single finger bone from Denisova Cave in Siberia was — and is — an astonishing achievement, both for the discovery itself and for its implications (D. Reich *et al. Nature* **468**, 1053–1060; 2010). It reinforces hints that the scarcity of human fossils belies what might once have been hitherto unimaginable diversity. The finding, reported in the same paper, that Denisovan DNA lives on in people from southeast Asia and the western Pacific, just as Neanderthal DNA survives in Eurasians generally, proves that fossils tell us much less than we would like of the human career. And as with Iwo Eleru, so with DNA: there are signs that the genomes of some modern Africans contain elements derived from archaic hominins not found in the fossil record (M. F. Hammer *et al. Proc. Natl Acad. Sci. USA* **108**, 15123–15128; 2011).

These early human relatives left signs of their passing as evanescent and enigmatic as the Cheshire Cat from *Alice's Adventures in Wonderland* — slowly fading from view, with just its smile hanging on, until that, too, disappears. ■



# Energy hit

*Germany's decision to slow the expansion of green-energy production is a reasonable move.*

For a few hours around lunchtime on a bright and windy day last month, Germany paid people to use electricity. The country's investment in renewable energy has paid off so handsomely that green power sources produced almost enough electricity to meet national demand. Because the rest of the energy infrastructure — nuclear, coal and gas stations — was online too, the temporary surplus pushed prices on Germany's spot market on 8 May into the negative. For a brief time, the more electricity that commercial customers used, the more money they made.

Just as the introduction of negative interest rates — banks charging customers to deposit money — have prompted more questions about the state of the world's economy, so the events of 8 May have been used to throw stones at Germany's energy transition, the *Energiewende*.

Critics of the *Energiewende* blame the generous support for renewables — including a guaranteed above-market price for producers, and grid priority for wind and solar power — for such grotesque market distortions and for Germany's relatively high household electricity bills. Indeed, private consumers in the country pay more than €20 billion (US\$23 billion) in annual surcharges for the fixed feed-in tariffs that go to individual producers. In response to criticism, the government last week agreed to slow the scale and pace at which Germany will further expand renewable energy over the next decade.

Electricity that is produced from renewable sources (including hydropower) has tripled in Germany over the past decade, and now provides (on most days) almost one-third of domestic electricity generation. The 19% growth in renewable-power generation last year was the largest in at least a decade. In terms of renewable-power capacity, Germany is third in the world, behind China and the United States (see *Trendwatch*, page 157). But at 1.1 kilowatts per capita, the 92 gigawatts that the nation produced in 2015 represents more than twice the renewable power per capita of any other large economy.

The planned amendment of Germany's renewable-energy act, agreed

in principle last week, sets a 45% cap on the amount of renewable electricity generation by 2025. And, as demanded by Brussels, future promotion of wind and solar energy will be linked to tenders, favouring producers who generate renewable power at the lowest price. Unsurprisingly, the plan has upset parts of the renewable-energy sector — in particular, smaller companies, and millions of homeowners who have

**“What Germany's case illustrates is that energy transition must be backed by comprehensive plans.”**

invested in lucrative rooftop solar panels — that fear the loss of market opportunities and revenue. Meanwhile, green lobbyists say the reforms are a concession to the fossil-fuel sector that could send a false signal to countries that regard Germany as a role model. Overall, it might seem that the *Energiewende* — and, by extension, renewables investment elsewhere — has failed, or is falling from political grace in troubled times.

That is wrong. Opinions will rightly differ on which fiscal and policy measures are best suited to promoting renewables. But controversies of microeconomics will not derail the grand decarbonization project that is under way.

What Germany's case does illustrate is that this transition must be backed by comprehensive energy plans. To incorporate fast-growing, decentralized power generation into electricity grids requires improved networks, reliable tools to predict supply and demand, efficient storage and more-flexible conventional plants.

Science can help. In April, the German federal government approved a 10-year, €400-million programme for research on the technology — smart grid technology and energy storage, for instance — that will be needed for an energy system and a market dominated by renewable sources. Funders should ensure that projects address the problems faced by suppliers and users of electricity under real-world market conditions. Sceptics may celebrate the wrinkles as renewables bed into the energy market, but the long-term trend has been set.

Of course, Germany and other countries will continue to depend on coal and gas to meet electricity demands for at least a couple of decades — but then, massive subsidies (\$500 billion globally in 2014, against \$135 billion for renewables) mean that fossil-fuel power continues to be offered at a knock-down price. A realistic, climate-friendly energy system for the future demands greater reform of the oil and coal markets, rather than changes to the renewables sector. ■

# Second chances

*The line between compliance and misconduct is finer than you might think.*

There are many reasons why a research paper could be retracted, although the one that tends to dominate public discussion is research misconduct.

Similarly, there are many different definitions of research misconduct, but the one that tends to draw attention is deliberate deception and data fraud. That can help to explain why, when *Nature* ran a news story in January 2013 about a new course that would attempt to rehabilitate misconduct offenders, many of the online comments below the story were negative (see *Nature* 493, 147; 2013).

One response was typical: “If a scientist wilfully pollutes the scientific record with damaging, self-serving, purposeful lies, it is too late. If they have received funding for that fraud, I am still mystified as to why the granting agency does not seek to recoup their costs in court. Now, someone suggests we further train them on something that is obvious to any scientist with half a brain on their shoulders?”

More than three years on, the architects of the rehab course offer

a progress report, which appears as a Comment article on page 173. More than three dozen researchers have been through its doors, and, according to the authors, most leave as better scientists than when they arrived.

Attendees do not include the high-profile data fraudsters whose offences are so serious that they are fired. By definition, researchers on the course are scientists who have been caught out but whom institutions want to keep.

Most of them had seen their research privileges suspended — with offences ranging from plagiarism and poor oversight to falling foul of the rules and regulations on animal welfare and informed consent. Despite the ‘research misconduct’ label, instances of conscious wrong-doing were rare. As one participant said: “Prior to this situation, I tried to follow the spirit of the law. Now I try to follow the letter of the law.”

Two points stand out. First, the typical character and personality of these scientists, and their knowledge and attitudes, were no different from those of you and your colleagues. Misconduct, the authors say, can be down to circumstance: “we believe that most researchers may be susceptible”. And second, those circumstances are becoming more common.

The most common cause of an offence was a lack of attention, prompted, among other things, by being too busy and trying to juggle too many projects. Sound like anyone you know? ■





## Gene editing can drive science to openness

*The fast-moving field of gene-drive research provides an opportunity to rewrite the rules of the science, says Kevin Esvelt.*

**T**he emergence of gene-drive systems — which spread engineered mutations quickly through populations — means that a single released organism could eventually alter most of its local population, and quite possibly all populations of the species throughout the world. Any accidental release, even if there was no ecological damage, would surely damage public trust and prompt harsh restrictions on research.

The US National Academy of Sciences released guidelines this week for the responsible conduct of gene-drive research. The report comes almost two years after the first published description of how the CRISPR–Cas9 genome-editing technology could enable gene drives in many different organisms. That's a fast turnaround for the academy, but an eternity for the field: in that time, scientists have demonstrated CRISPR-based gene-drive systems in four species.

The report makes some sensible suggestions, such as phased testing and ecological-risk assessments, but if we're going to develop proper safeguards for gene drives or other powerful technologies, we need to fix a greater problem: the closed-door nature of science.

No one would rationally design the current scientific enterprise. It is wasteful and inefficient. Researchers repeatedly run into the same problems and unknowingly duplicate efforts. It stunts collaboration: we never learn who has the other piece of a puzzle unless we run into them at a conference. It wastes time on endless grant-writing. It's terrible for researcher well-being: competitive pressure ruins playful discovery and creation.

And it's unsafe. Regulation will always be too slow. Science is too vast for researchers to reliably foresee the consequences of their work. The problem was neatly summarized by atom-bomb pioneer Robert Oppenheimer: "When you see something that is technically sweet, you go ahead and do it, and you argue about what to do about it only after you have had your technical success."

Some technical successes are not to be pursued. But others are desperately needed. How can we hope to tell the difference when science is done behind closed doors?

There are signs of progress. My colleagues and I publicly discussed the probable consequences of a CRISPR-based gene drive before doing any experiments. And many gene-drive researchers have already worked together to improve safety and call for transparency. But this has been done on an informal basis. For example, my group saw a gene-drive paper by another laboratory and was able to suggest changes — the need for extra safeguards to prevent an accidental release — but only because we received an in-press copy of the publication from a journalist.

Sadly, open and responsive science flies in the face of current incentives. Scientists who disclose their ideas are often 'rewarded' by being

scooped by another lab, rather than by being recognized for their creativity. It is a prisoner's dilemma. The benefits come from cooperation by everyone. But by participating you risk being exploited by people who steal your idea, get it working before you do, and claim the credit.

Gene-drive research offers a way out. The field is new and small, and many of us have already worked together to publish a joint recommendation calling for future experiments to use multiple stringent confinement strategies. Several groups already disclose proposed and ongoing gene-drive research and invite feedback, and active discussions between researchers and funders seek ways to ensure that everyone will be similarly forthcoming.

My group and others will soon launch the Responsive Science Project to enable gene-drive scientists to share their plans and research with one another and with interested communities. We hope that it will become

a central repository of ideas and information relevant to gene-drive research that will permit open assessment and critique before experiments begin.

Journals could help by offering incentives to persuade scientists to share their proposals. When a paper is published by authors who didn't play by the new rules (to share what they're doing and collaborate with the people who first shared the key ideas), journals could check the repository to identify scientists who deserve a share of the credit and invite them to write an accompanying piece. Similarly, all funders should require immediate public disclosure of proposals involving gene drives, as well as regular public updates on the status of funded research.

If this attempt at open science works for one field, it could expand to encompass research on

other shared-impact technologies and to fields beyond. That alone is reason enough to try the approach. But gene-drive technology is also unique in that its very nature demands a new approach.

Because the consequences of mistakes involving gene-drive organisms could affect communities outside the laboratory, scientists have an obligation to openly share their plans, invite suggestions and concerns, disclose experimental results as soon as possible, and redesign the technology as needed. Applied to gene drives, such an approach will also have a greater chance of earning popular support for applications that could save millions of human lives and rescue numerous species from extinction.

We should ensure that gene-drive research is open and responsive — then drive those changes through the scientific ecosystem. ■

**Kevin Esvelt** is leader of the Sculpting Evolution group at the MIT Media Lab, Massachusetts Institute of Technology, Cambridge, Massachusetts.  
e-mail: [esvelt@mit.edu](mailto:esvelt@mit.edu)

**OPEN  
AND RESPONSIVE  
SCIENCE  
FLIES IN THE  
FACE  
OF CURRENT  
INCENTIVES.**



# RESEARCH HIGHLIGHTS

Selections from the  
scientific literature

## MOLECULAR BIOLOGY

### CRISPR tweaked to edit RNA

The CRISPR–Cas9 gene-editing system snips DNA, but a newly characterized version targets RNA instead.

The CRISPR–Cas system is used by many bacteria to combat viruses. Feng Zhang of the Broad Institute of MIT and Harvard in Cambridge, Massachusetts, Eugene Koonin of the US National Institutes of Health in Bethesda, Maryland, and their colleagues mined this natural diversity for alternatives to the DNA-cutting Cas9 enzyme. They found that an enzyme called C2c2 from the bacterium *Leptotrichia shahii* can be programmed to cut specific, single-stranded RNA targets in another bacterium, *Escherichia coli*.

With further tweaks, the system could be used to attach fluorescent tags to RNA, direct RNA to specific compartments in the cell or otherwise chemically modify RNAs to study their function. *Science* <http://dx.doi.org/10.1126/science.aaf5573> (2016)

## GEOLOGY

### Magma pool under New Zealand

Molten rock is accumulating in a magma chamber beneath New Zealand, raising questions about volcanic hazards.

Ian Hamling and his colleagues at GNS Science in Lower Hutt, New Zealand, used satellite radar data to study ground motions in the Taupo Volcanic Zone, an area of high volcanic activity. They found one region, adjacent to this area, where the ground rose by around 5 millimetres per year from



## GLACIOLOGY

### Early signs of ice retreat

Two studies show that Antarctica has been losing ice for longer than previously thought.

A team led by Shujie Wang at the University of Cincinnati in Ohio studied recently declassified images taken by US spy satellites. They found that glaciers feeding the Antarctic Peninsula's Larsen B Ice Shelf (pictured in 2000) were already accelerating towards the sea between 1963–79 and 1979–86, long before the shelf's spectacular collapse in 2002.

In separate work, Frazer Christie at the University of Edinburgh, UK, and his colleagues used scientific satellites to confirm that ice has been retreating along the coast of West Antarctica since at least 1975.

*Geophys. Res. Lett.* <http://doi.org/bjm3>; <http://doi.org/bjm4> (2016)

the 1950s onwards. That rate more than doubled to about 12 millimetres a year in the mid-2000s, and has since dropped back to the lower rate. Calculations suggest that about 9 million cubic metres of magma pushed its way into the crust each year during peak growth, about 10 kilometres below the surface.

It's not clear whether the magma chamber will increase

the risk of volcanic eruptions. *Sci. Adv.* 2, e1600288 (2016)  
For more on this research, see [go.nature.com/28ew7kh](http://go.nature.com/28ew7kh)

## AGEING

### Chemical extends worm lifespan

A chemical lengthens the nematode worm's lifespan by interfering with its perception

of whether food is present.

Model organisms are known to live longer when they are fed a restricted diet. Mark Lucanic and Gordon Lithgow at the Buck Institute for Research on Aging in Novato, California, and their colleagues screened 30,000 synthetic compounds and found several that extended the lifespan of the nematode *Caenorhabditis elegans*. The most potent, NP1, mimicked the effects of dietary restriction by masking the activity of a sensory pathway that normally signals that food is abundant. The chemical does this by boosting signalling of a specific neurotransmitter called glutamate to the pharynx, which in nematodes is a tube-like organ that pumps food into the gut.

Further investigation of nutrient-sensing pathways could identify other life-extending chemicals, the authors say.

*Aging Cell* <http://doi.org/bjhh> (2016)

## ASTROPHYSICS

### Relativity passes black-hole test

General relativity holds true, even under the extreme conditions of colliding black holes.

In 2015, the Advanced Laser Interferometer Gravitational-Wave Observatory (LIGO) saw the first evidence of gravitational waves, which had been created by two merging black holes. Walter Del Pozzo at the University of Birmingham, UK, and his colleagues on the LIGO collaboration and its European partner, the Virgo collaboration, compared the signal with those predicted by simulations based on general relativity. The teams found that the observations matched the predictions to a high degree, as they had in previous tests

LANDSAT 7 SCIENCE TEAM/GSFC/NASA



TADD TRUSCOTT

under much weaker gravitational fields. This was the first direct test of Einstein's theory of general relativity under such extreme space-time warping and fast-moving conditions.

With planned boosts to the LIGO detectors' sensitivity, future observations could be used to test other theories of gravity and hypothesized alternatives to black holes, say the authors.

**Phys. Rev. Lett.** 116, 221101 (2016)

## MARINE SCIENCE

## Plastic pollution hurts perch

Tiny fragments of plastic in the ocean could change fish behaviour and decrease their survival.

Research indicates that the world's oceans are polluted with many thousands of tonnes of 'microplastic' debris — particles measuring less than 5 millimetres in diameter. Oona Lönnstedt and Peter Eklöv at Uppsala University in Sweden exposed European perch (*Perca fluviatilis*) to levels of microplastic similar to those found in the environment. Although 96% of fertilized eggs not exposed to plastic hatched, only 81% of those placed in water with high levels did so. Moreover, 46% of fish larvae that had been raised in a tank containing plastic-free water were still alive after 24 hours in a tank with a predatory pike, whereas 100% of those raised with high levels of plastic were eaten within 16 hours.

Larvae reared with high concentrations of plastic did not show anti-predator responses such as freezing or reduced movement when exposed to alarm signals from other animals.

**Science** 352, 1213–1216 (2016)



## GENOMICS

## Ancient dog DNA shows dual origins

The first complete genome sequence of an ancient dog suggests that dogs were independently domesticated twice, in two different regions.

Researchers have debated whether domestic dogs originated in Asia or Europe about 15,000 to 12,500 years ago. Laurent Frantz of the University of Oxford, UK, and his team sequenced the mitochondrial DNA of 59 ancient dogs and the complete genome of a 4,800-year-old dog from Ireland. They also analysed DNA from hundreds of modern dogs and wolves, and found that populations of Western European and East Asian canines diverged several millennia after the first appearance of the animals.

Owing to a lack of archaeological evidence of ancient dogs between these regions, the authors propose that the animals were domesticated separately in Western Europe and East Asia from distinct wolf populations. **Science** 352, 1228–1231 (2016) For more on this research, see [go.nature.com/1pzqqvr](http://go.nature.com/1pzqqvr)

## EVOLUTION

## Age robs monkeys of vocal control

Monkeys lose the ability to consciously control their calls as they age, which may have limited the evolution of language in non-human primates.

Steffen Hage and his colleagues at the University of Tübingen in Germany studied the vocalizations of two male captive rhesus macaques (*Macaca mulatta*; pictured) over a roughly five-year period. The monkeys were trained to produce a specific call in response to a coloured cue to receive a reward. At five



years old, the macaques scored highly, but by eight years of age neither monkey could perform the task. Adult macaques still produced spontaneous, instinctive calls in their enclosure, indicating that they maintained vocal ability.

Language may have evolved in humans by first extending the vocal flexibility of juveniles into adulthood.

**J. Exp. Biol.** 219, 1744–1749 (2016)

## BOTANY

## How desert moss drinks from air

Researchers have revealed minuscule features on the leaves of a common desert plant that allow it to collect water from moist air.

*Syntrichia caninervis* (pictured) is a small moss that lacks roots. To understand how it uses its leaves to capture moisture, Tadd Truscott of Utah State University in Logan and his co-workers altered the relative humidity in their lab and used high-speed cameras and electron microscopy to study the plant's response. They found that a hair-like structure called an awn at the tip of each leaf has grooves and barbs that collect and transport water. Nanogrooves about 200 nanometres wide allow water from humid air to nucleate on the awn's surface, and larger microgrooves collect bigger water droplets from fog.

Small barbs along the cone-shaped awn provide places for

droplets to collect before being transported down the awn to the leaf.

**Nature Plants** <http://dx.doi.org/10.1038/nplants.2016.76> (2016)

## INFECTION

## Bacterium could curb malaria

West African mosquitoes infected with the bacterium *Wolbachia* are less likely than uninfected ones to carry the malaria parasite *Plasmodium*.

*Wolbachia* infection has long been proposed as a way to reduce the spread of mosquito-borne diseases such as malaria. To study natural *Wolbachia* infection, Flaminia Catteruccia at the Harvard T. H. Chan School of Public Health in Boston, Massachusetts, and her colleagues collected and studied 221 *Anopheles coluzzii* mosquitos from a village in Burkina Faso. They found that about half of the insects carried a *Wolbachia* strain. Only one infected mosquito (less than 1%) was also infected with *Plasmodium*, whereas roughly 10% of the 105 mosquitos free of *Wolbachia* tested positive for the malaria parasite.

Mathematical modelling suggested that even at this rate of *Wolbachia* infection, the bacterium could decrease the prevalence of malaria in humans.

**Nature Commun.** 7, 11772 (2016)

### NATURE.COM

For the latest research published by Nature visit:

[www.nature.com/latestresearch](http://www.nature.com/latestresearch)

JOEL SARTORE/NATL GEOGR/GETTY



# SEVEN DAYS

The news in brief

## POLICY

### Forensics guidelines

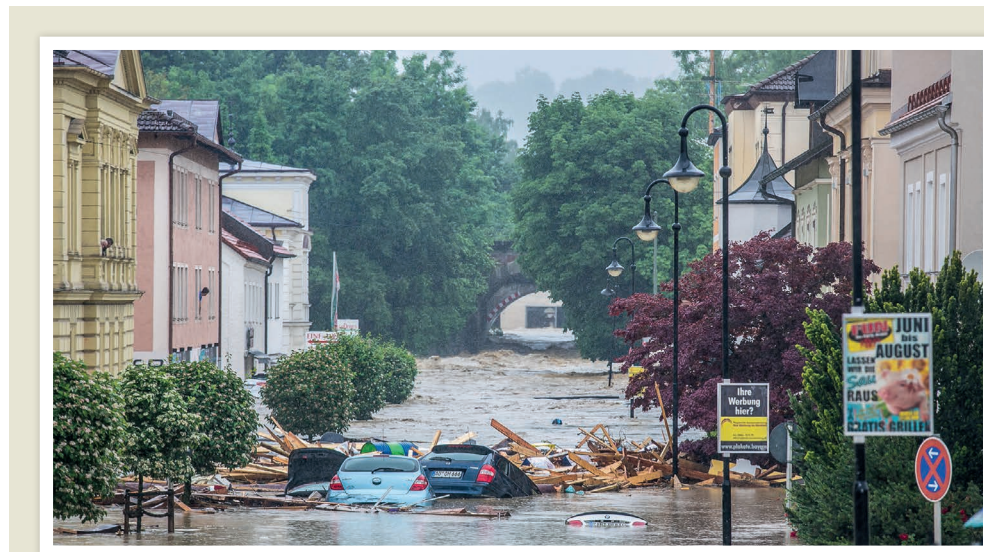
On 3 June, for the first time, the US Justice Department issued recommendations meant to help to guide expert testimony and lab reports in forensic science. If adopted, they would apply to all agencies that issue forensic reports and would cover seven disciplines, including body-fluid testing, fingerprints and toxicology. The department will issue further guidelines later in the summer regarding genetic testing and several other forms of forensic evidence.

### Policy labours lost

An independent inquiry published on 2 June found that most UK government departments don't keep track of the policy-linked research that they have paid for in areas such as health, education or climate change. Only 4 out of 24 government departments contacted through freedom-of-information requests keep a centralized database

### SEEKING GREAT MENTORS!

For more than ten years, *Nature* has been recognizing outstanding scientific mentors around the world. This year's awards for great mentors are focused on the US states of Washington, Oregon and California. Each prize is an award of US\$10,000. For details of the competition and guidance about nominating candidates for lifetime and mid-career awards, see [www.nature.com/nature/mentoringawards/uswestcoast](http://www.nature.com/nature/mentoringawards/uswestcoast). Nominations must be submitted by mentees of the nominees, and the closing date is 8 August 2016.



DANIEL SCHARINGER/AFP/GETTY

## Extreme weather rocks Europe

Severe storms and heavy rain caused havoc last week in parts of Western and Central Europe. On 1 June, a flash flood triggered by extreme downpours devastated the town of Simbach am Inn in southern Germany, killing 7 people and causing damage across the region in excess of €1 billion (US\$1.1 billion). In Paris on 3 June, the River Seine's rising floodwaters led authorities to close the Louvre

and Orsay museums and move artworks to safety. And on 4 June, Germany's largest rock festival, near Mendig, was aborted after lightning struck at least 80 visitors. Because warmer air holds more moisture, rising temperatures are increasing the odds of intense precipitation. Worldwide, record-breaking rainfall events have been on the rise over the past 30 years.

of commissioned research, according to Sense About Science, the London-based science-advocacy group that commissioned the inquiry. The report also points to several cases in which the publication of sensitive findings — including on drug use and immigration — was delayed owing to political concerns. See page 164 [go.nature.com/lsqhudo](http://go.nature.com/lsqhudo) for more.

## BUSINESS

### Novartis Shanghai

The pharmaceutical giant Novartis opened an extensive campus in Shanghai, China, on 2 June. The campus hosts the third major site of the Novartis Institutes for

Biomedical Research (NIBR), where around 300 scientists will use epigenetic approaches to develop treatments for cancers and other diseases that are prevalent in China and Asia. The other two main NIBR campuses are in Cambridge, Massachusetts, where research includes cancer, cardiovascular disease and diabetes; and in Basel, Switzerland, which focuses on autoimmunity, transplantation and inflammation.

### Solar power in Cuba

A British solar company, Hive Energy, has secured a contract to build the first utility-scale solar power plant in Cuba. On 3 June, the company, based near Southampton, announced plans to complete

a 50-megawatt solar-power project by 2018 in a free-trade zone in the Cuban port city of Mariel. The project is one of many that the Cuban government committed to pursuing under the 2015 Paris climate agreement.

## RESEARCH

### Stem-cell setback

The biopharmaceutical company StemCells of Newark, California, has prematurely terminated a phase II clinical trial examining whether cells derived from fetal brains could help to fix damaged spinal cords and increase muscle strength in paralysed limbs. Interim analysis of the 17 people treated so

far showed no significant improvement, even though experiments in mice with spinal damage had been positive. The company is planning to close down its operations. Other trials of stem-cell preparations to treat spinal-cord injury are continuing.

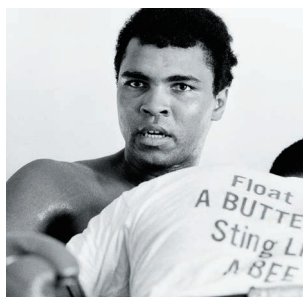
## Shared data

The US National Cancer Institute (NCI) unveiled a platform for sharing genomic and clinical data on 6 June at the annual meeting of the American Society of Clinical Oncology in Chicago, Illinois. The platform, called the Genomic Data Commons, will be a key component of US vice-president Joe Biden's National Cancer Moonshot and President Barack Obama's Precision Medicine Initiative.

## PEOPLE

### Tribute to Ali

Millions of admirers, including US President Barack Obama, are mourning the death of Muhammad Ali (pictured) — legendary boxer, political activist and, in his later years, advocate for Parkinson's disease research. The three-time heavyweight champion died on 3 June, aged 74, having lived for more than 3 decades with the degenerative brain disease. In 1997, Ali co-founded the



Muhammad Ali Parkinson Center, a research and treatment facility in Phoenix, Arizona.

## AWARDS

### Kavli Prize

On 2 June, nine scientists were awarded the 2016 Kavli Prize, a biennial award that honours individuals for seminal advances in astrophysics, nanoscience and neuroscience. This year's laureates come from Germany, Switzerland, Britain and the United States. The co-founders of the Laser Interferometer Gravitational-wave Observatory (LIGO), Ronald Drever, Kip Thorne and Rainer Weiss, won for the detection of gravitational waves. Gerd Binnig, Christoph Gerber and Calvin Quate were recognized for their development of atomic-force microscopy. Eve Marder, Michael Merzenich and Carla Shatz won for the discovery of key brain mechanisms.

Recipients from each category shared a US\$1-million cash prize.

### Courage to think

Egyptian students and scholars wrongfully detained in their home country are being honoured on 9 June with the 'Courage to Think Defender Award' by academic-freedom advocates Scholars at Risk. The group cites an "overwhelming crackdown" in recent years on Egypt's higher-education community, including "reported use of violence, wrongful prosecutions and imprisonment, professional retaliation and travel restrictions against scholars and students across the country". According to data obtained by the Egyptian human-rights group Association for Free Thought and Expression, more than 2,000 university students and professors in Egypt have been detained by security forces since July 2013.

## EVENTS

### Campus murder

On 1 June, an armed 38-year-old man entered the engineering department at the University of California, Los Angeles (UCLA), and fatally shot William Klug, a mechanical and aerospace engineering professor. Police think that the killer, who

received his doctorate with Klug's guidance in 2013, then shot himself. The authorities have not confirmed a motive for the murder-suicide, but stated that the gunman had accused Klug of stealing his computer code and giving it to someone else. Investigators found a 'kill list' in the gunman's Minnesota home that included Klug, another UCLA professor (who was unharmed) and a woman whom police subsequently found dead in a Minnesota suburb.

### Pure gravity

The European Space Agency's LISA Pathfinder team announced on 7 June that its mission to test technology for a space-based gravitational-wave detector has been a success. Results from the first two months of operation showed that two metal cubes housed in the spacecraft were under free-fall through space, affected almost solely by gravity, with perturbations kept five times lower than required. A full-scale observatory, which will use similar cubes in spacecraft placed millions of kilometres apart to measure passing gravitational waves, is scheduled for launch in 2034.

### Brazilian anger

Brazil's science ministry will not be reinstated, despite researchers' protests, the country's government says. The ministry was merged with the telecommunications ministry by interim President Michel Temer, who took office in May after the impeachment of Dilma Rousseff (see *Nature* 533, 301; 2016). In the past week, scientists — already angry at massive budget cuts — have taken to the streets in Natal, Rio de Janeiro and São Paulo to protest against the merger. But Gilberto Kassab, who heads the science-telecomms 'superministry', has told *Nature* that the merger will be maintained.

➔ **NATURE.COM**

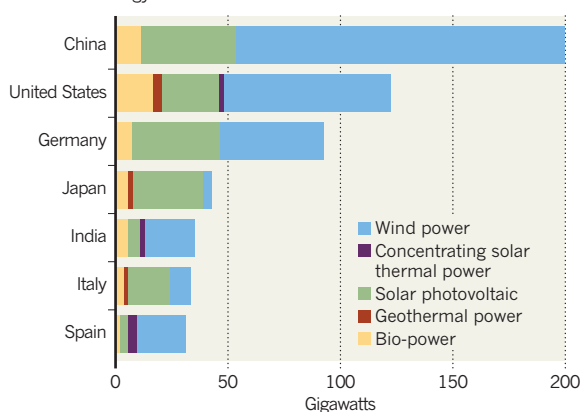
For daily news updates see:  
[www.nature.com/news](http://www.nature.com/news)

## TREND WATCH

Global power-generating capacity from renewable sources excluding hydropower had climbed to 785 gigawatts by the end of 2015, after a record 120 gigawatts were added during the year. China, which hosts more than 25% of the world's non-hydro renewable capacity, is the top country in absolute terms, followed by the United States and Germany. China and the United States both lag far behind Germany in per capita capacity. Including hydropower, renewables have a 24% share in world energy.

### RENEWABLE BOOM

Wind power and solar photovoltaics dominate the fast-growing renewable-energy sector.





# NEWS IN FOCUS

**MEDICINE** A fairer way to distribute untested drugs **p.160**

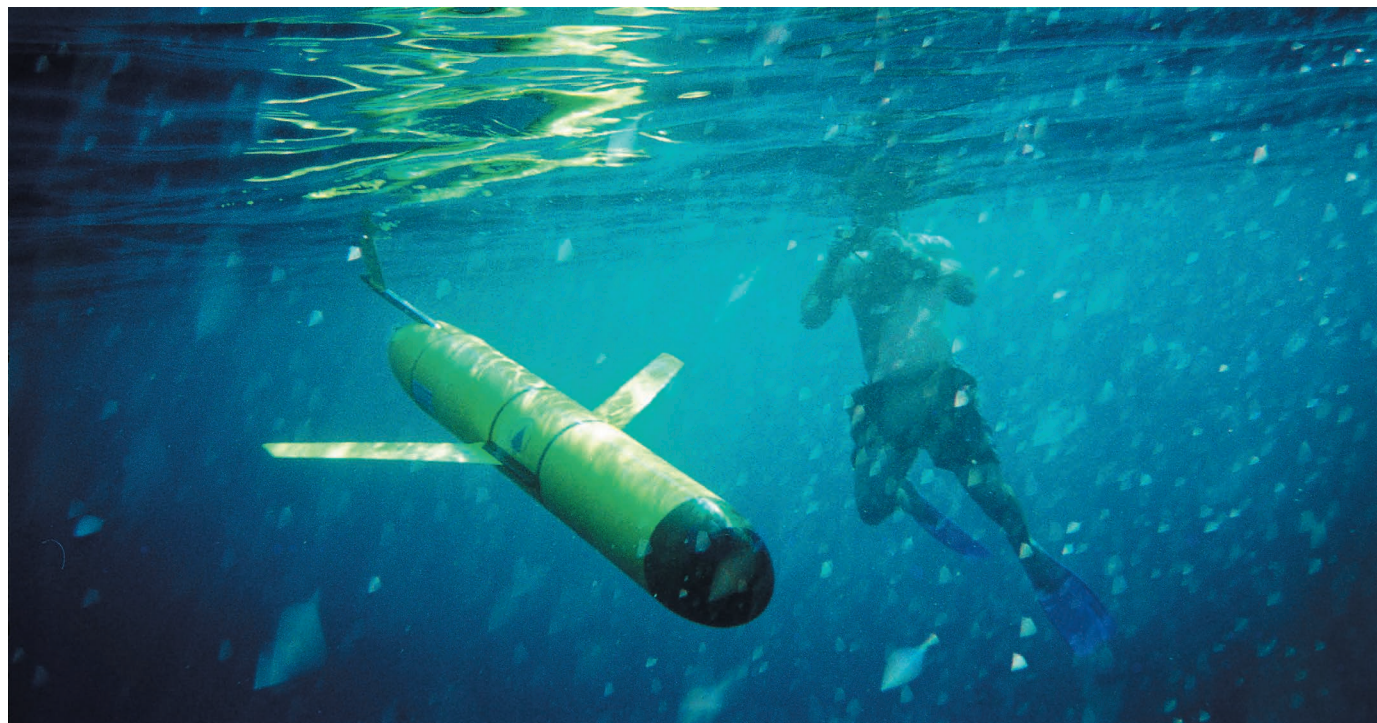
**HEALTH** Peru's gold-mining boom drives mercury contamination **p.162**

**GENOMICS** Plan to 'write' human genome triggers mixed reaction **p.163**



**MICROBIOLOGY** Pieter Dorrestein eavesdrops on chemical conversations **p.170**

PAUL FUCILE/WHOI



Gliders are among the instruments used at the seven monitoring sites of the US Ocean Observatories Initiative.

## OCEANOGRAPHY

# US ocean-observing project launches at last

*Data from a network of deep-water observatories are now streaming in real time.*

BY ALEXANDRA WITZE

Nearly ten years, US\$386 million and many grey hairs after it got the go-ahead, an enormous US ocean-observing network is finally up and running.

On 6 June, the National Science Foundation (NSF) announced that most data are now flowing in real time from the Ocean Observatories Initiative (OOI), a collection of seven instrumented arrays. Oceanographers have the chance to test whether the technologically complex and scientifically unprecedented project will ultimately be worth it.

"It has been stressful," says Richard Murray, the NSF's director for ocean sciences. "It's not for the faint-hearted."

The raw-data streams came online in April — months behind schedule, in part because of a 2014 switch between university subcontractors.

Through an open-records request, *Nature* obtained more than 1,200 pages of e-mails between project managers at the NSF and the Consortium for Ocean Leadership in Washington DC, which built the observatory. The records reveal an extraordinary level of tension throughout 2014 and early 2015, as the final instruments were installed and the contract for

handling the data streams was switched from the University of California, San Diego, to Rutgers University in New Brunswick, New Jersey.

"Please excuse my display of stress in this email, but the InBox is overflowing with high-priority, short-fuse items — none of which deserve to be ignored — but all of which cannot be completed within the requested time frames," Timothy Cowles, then programme director at the Consortium for Ocean Leadership, wrote to the NSF in January 2014.

The NSF cited cost overruns and performance delays as justification for changing the cyberinfrastructure contract later that

► year. In April 2015, an underwater volcano laden with OOI instruments erupted, just as scientists had predicted — but the live data had not yet flowing to the wider scientific community.

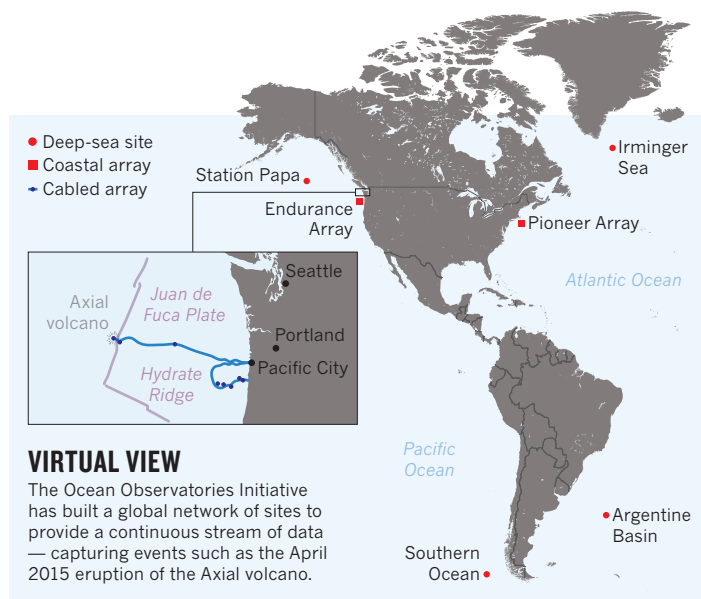
Now, about 85% of OOI data are available in real time on the project's website, with the percentage growing every week, says Greg Ulses, the current programme director at the Consortium for Ocean Leadership. The information — on factors such as temperature and salinity — streams from more than 900 sensors at the 7 sites.

The OOI consists of one high-tech cable on the tectonically active sea floor of the northeast Pacific Ocean, together with two lines of oceanographic instruments — one off the US east coast and the other off the west coast — and four high-latitude sites, near Greenland, Alaska, Argentina and Chile. Each array involves a combination of instruments, from basic salinity sensors to sophisticated underwater gliders.

The NSF built the network as a community resource, hoping to stimulate an era of virtual oceanography in which scientists explore real-time data sets open to all (see 'Virtual view').

"We know the data are valuable," says Lisa Campbell, a biological oceanographer at Texas A&M University in College Station. "How to implement it is what we're working on."

Those involved in the OOI's painful birth are happy to see it working at last. "When I finally got through and saw the real-time data, I shouted so loud someone had to come down the hall and close the door," says Glen



Gawarkiewicz, a physical oceanographer at the Woods Hole Oceanographic Institution in Massachusetts.

The array off the coast of Massachusetts has already captured some unprecedented observations, he says. In 2014, it measured air-sea fluxes when a hurricane passed overhead. The following winter, it measured dramatic shifts in the boundary at which shallow waters interact with deep ones. "That has tremendous practical implications, because there's a lot of commercial fishing in that area," Gawarkiewicz says. Using OOI data, he is now working with local fishers to share real-time information on changes in temperature and currents.

The west-coast array has studied a warm blob of water linked to weather patterns that are strengthening the ongoing drought in California. And in the North Atlantic, off the coast of Greenland, OOI scientists have coordinated their measurements with those of

others, such as an international programme to measure heat flow in this key region. "These are high-scientific-value sites that we have dreamed about, and now we have occupied them," says Robert Weller, a physical oceanographer at the Woods Hole Oceanographic Institution.

But the OOI's future remains murky. A 2015 review of US ocean-science priorities suggested that the programme's operational budget should be slashed by 20%, to around \$44 million a year. Yet each of the arrays must be serviced every year or two to replace broken instruments and install new ones.

The NSF has not yet decided how it will save that 20%.

Later this year, the agency will solicit bids to manage the OOI for the next five to ten years. Who responds, and with what suggestions, will help to determine what gets cut. "We built this thing, and will be funding operations for what the community feels is best," says Murray.

Ultimately, there is no metric for what constitutes a successful OOI. Ulses says that the project needs to run for a full year before managers can assess which scientists are using which data, and how stable and successful the data streams are.

Weller would like to see a set of OOI measurements become as iconic as the records of atmospheric carbon dioxide levels taken at Mauna Loa, Hawaii, since the 1950s. "On any given day, I step back," he says, "and am still sort of amazed that it's all out in the water and most of it's working." ■

## BIOMEDICINE

# Panel tackles 'compassionate use'

*Companies pressured by social-media appeals seek fair way to allocate last-ditch treatments.*

BY SARA REARDON

Nancy Goodman wanted to spend as much time as possible with her dying child. But even as ten-year-old Jacob's brain cancer worsened, Goodman spent months contacting pharmaceutical companies that were developing drugs that might help him.

'Compassionate-use' laws in the United States allow pharmaceutical companies to provide unapproved drugs to patients in desperate need, but many firms provide little or no information on how to request these treatments. They are often reluctant to supply drugs

in response to such pleas, especially if drug stocks are limited, although media campaigns on behalf of individual patients can sometimes embarrass firms into providing unapproved treatments. Anecdotes suggest that money and connections are also influential.

Now, ethicists and medical experts are testing what they hope is a fairer system to distribute drugs in short supply. The approach, presented on 6 June at the American Society of Clinical Oncology meeting in Chicago, Illinois, is inspired by the method used to prioritize organ transplants. In a test case, researchers worked with Janssen Pharmaceuticals to

determine how to distribute limited supplies of daratumumab, an experimental drug intended to treat multiple myeloma.

The 10-person panel combed through 76 anonymized applications to determine how likely the drug was to work for each person, ultimately approving 60. "It's hard to say no, because people die," says Arthur Caplan, a bioethicist at New York University's Langone Medical Center who is leading the effort. But he says that a systematic approach could help companies to make unbiased decisions.

In Goodman's case, six of the eight companies that she contacted never responded. The



other two declined to give her son their drugs because the treatments had never been tested in children. Jacob died in 2009, and his mother went on to found the advocacy group Kids v Cancer in Washington DC.

There are many legitimate reasons that companies might refuse to provide unapproved drugs, says Aaron Kesselheim, who studies health-care ethics at Brigham & Women's Hospital in Boston, Massachusetts. People who request such treatments are often very ill, and companies worry that their deaths while receiving the drug would reduce the compound's chances of approval from the US Food and Drug Administration (FDA). Giving patients access to experimental drugs could also discourage them from enrolling in controlled trials that might assign a placebo, and would leave less drug available for use in the trial.

"These requests are some of the most difficult decisions I face as a physician," says Amrit Ray, chief medical officer of Janssen in Titusville, New Jersey. "It's a trade-off we have to consider carefully."

Since 2014, 28 US states have enacted 'right-to-try' laws, which allow companies to provide drugs to patients without involving regulators. Caplan calls these "feel-good" laws, because the FDA approves most of the compassionate-use requests that it receives. (It is not clear how many applications are denied by companies and never reach the FDA.)

Vickie Buenger, president of the advocacy group Coalition Against Childhood Cancer in Philadelphia, Pennsylvania, says that right-to-try statutes contribute to patients' misunderstanding about the factors that go into a decision to supply or deny access to a drug. "It implies that companies and the FDA are either angels of mercy if they come through, or devils who have no compassion if they withhold it."

This lack of clarity, and poor communication by companies, has led many patients and their families to launch social-media campaigns to secure unapproved drugs.

Perhaps the most famous case came in 2014, when the family of seven-year-old Josh Hardy began a Facebook campaign for an unapproved antiviral drug called brincidofovir to treat a life-threatening infection. Its manufacturer, Chimerix of Durham, North Carolina, had declined, on the grounds that giving the drug to Josh — and any subsequent petitioners — would leave less of the compound



Josh Hardy received an experimental drug after his family launched a massive social-media campaign.

available for an ongoing clinical trial. Within days, the Facebook page and Twitter campaign #savejosh were featured on national television. Chimerix quickly created a small clinical trial with Josh as its first patient.

"Every single CEO woke up the next morning and said, 'Oh my gosh, that might happen to me,'" says Elena Gerasimov, who directs a programme at Kids v Cancer that helps

**"These requests are some of the most difficult decisions I face as a physician."**

parents of children with cancer to petition companies for drug access. (The FDA is attempting to make this process easier. On 2 June, it released new forms to simplify the filing of compassionate-use appeals.)

Former Chimerix chief executive Kenneth Moch says that dozens of companies have since enlisted him as an adviser on such issues. His advice is simple: every company should create a transparent system to handle compassionate-use requests, guided by the FDA. That is in line with the advice of the Biotechnology Innovation Organization, an industry group in Washington DC that encourages its members to develop clear policies to explain whether they provide expanded access and to help physicians to request drugs. "That's the least we can

do, to facilitate people being able to contact us," says Kay Holcombe, the group's senior vice-president for science policy.

Caplan and Ray plan to test their system on another treatment later this year — possibly a mental-health drug or a childhood vaccine. Caplan hopes that more companies will adopt the approach, and imagines eventually creating a compassionate-use consulting panel to aid small companies.

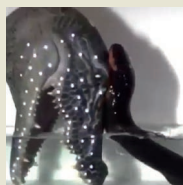
Moch cautions that the approach might not be appropriate for every drug or company, but he likes how it helps to level the playing field. "Had Josh been a 37-year-old guy who kicked his dog and smoked, he wouldn't have gotten the same support as a lovely seven-year-old boy," he says.

Patient advocates also support Caplan's system for distributing drugs. "Putting it in the hands of people who understand the drug's possibilities is a reasonable thing," Buenger says.

But many also want the FDA to create incentives for companies to provide drugs for compassionate use. Until that happens, or until companies adopt programmes such as Caplan's, social-media campaigns and other public appeals may be some patients' only option. "I'd do it," Goodman says. "I'd do anything to save my kid — anything to give Jacob a few more months." ■

## MORE ONLINE

### VIDEO



Electric eels leap from water tank to shock a fake alligator head [go.nature.com/eelsvideo](http://go.nature.com/eelsvideo)

### MORE NEWS

- CRISPR gene editing unleashed on RNA [go.nature.com/ghytrs](http://go.nature.com/ghytrs)
- 'Zombie volcano' grows beneath New Zealand [go.nature.com/ltt4aa](http://go.nature.com/ltt4aa)
- Ancient genomes suggest a dual origin for modern dogs [go.nature.com/rhxlsd](http://go.nature.com/rhxlsd)

### NATURE PODCAST



Researcher rehab; the hobbit's ancestry; and Google's quantum plans [nature.com/nature/podcast](http://nature.com/nature/podcast)



Many miners in Peru's Madre de Dios region use mercury to extract gold from sediments.

#### ENVIRONMENTAL SCIENCE

# Peru's gold rush raises health fears

*Gold-mining boom in southeastern Amazon is driving high levels of mercury contamination.*

BY BARBARA FRASER

Long-running concerns about the environmental effects of gold-mining in the Peruvian Amazon came to a head on 23 May. Peru's government declared a 60-day public-health emergency in an attempt to address the problem of mercury pollution caused by unregulated gold-mining along the Madre de Dios River.

Health-care and emergency workers are now providing medical and food aid for 25 affected villages, after a flurry of studies showed high levels of mercury in people, fish and sediments in the Madre de Dios region. The government estimates that some 48,000 people across 85,301 square kilometres have been affected.

"We now know with certainty what the source of the exposure is," says Peru's deputy health minister, Percy Minaya. "We are not going to solve this in two months, or even in a year, but the health ministry has to start." Symptoms of mercury poisoning include vomiting and diarrhoea. Extreme cases can lead to brain or kidney damage.

The Madre de Dios region has a long history of small-scale alluvial gold-mining, but the rise in international gold prices in the past decade has brought a boom in the activity.

Peru's National Institute of Statistics and Information in Lima reported on 23 May that gold production for March in Madre de Dios was 1,583 kilograms, up 28% on the same month last year.

The region's miners extract the gold by sluicing sediment to separate out gold-bearing sand, which they then mix with mercury to form an amalgamated lump of metal. Heating the lump vaporizes the mercury, leaving pure gold behind. The process sends an estimated 30–40 tonnes of mercury each year into waterways, where bacteria convert the metal into methylmercury. The methylmercury accumulates in fish, which are a key source of food for people in the Madre de Dios region.

Perhaps unsurprisingly, researchers have found high levels of mercury (above the maximum recommended by the World Health Organization) in hair samples from 40% of the Madre de Dios residents that they tested. The team, from Duke University in Durham, North Carolina, examined about 800 people who live along a major highway in the region, 100 people who live beside the river and 2,000 in the Amarakaeri Indigenous Reserve.

Some communities in the region are closer to the gold-mining activities than others, but the 40% exposure rate held across the highway,

river and reserve, says study leader William Pan, an epidemiologist with Duke's Global Health Initiative.

The presence of mercury in human hair usually indicates that a person has been exposed to the metal through a dietary source. Pan says that the Duke studies in Madre de Dios show a strong correlation between human mercury exposure and fish consumption. Since 2009, research by Pan's group (S. E. Diringer *et al. Environ. Sci. Processes Impacts* 17, 478–487; 2015) and by tropical ecologist Luis Fernandez at the Carnegie Institution for Science at Stanford University in California, have found high mercury levels in some species of fish, particularly in large catfish and fish that eat other fish.

#### BURDEN REDUCTION

Peru's government used the Duke team's latest study to decide which riverside communities should receive the emergency aid. Officials are trying to help affected residents to replace the high-risk fish in their diets with other sources of protein. During the emergency period, the government will give food, including canned ocean fish, and multivitamins to combat anaemia, to roughly 15,000 of the 48,000 people affected.

Pan says that these steps should reduce the body burden of mercury in people who also cut their consumption of contaminated fish, because the primary route of mercury exposure in the region seems to be through food.

The government is also considering whether its food aid should include supplies of the grain quinoa. Preliminary data from Duke's household surveys in the Madre de Dios region show a correlation between quinoa consumption and lower mercury levels.

Minaya says that the government's long-term plan also includes helping communities to establish fish farms. The emergency period is set to end days before a new president takes office on 28 July. But Minaya is confident that the next administration will continue to monitor and address the mercury pollution problem, despite opposition from regional and local government officials.

These officials have criticized the emergency decree, arguing that the link between people's mercury levels and fish consumption is not proven. The officials are also worried that the public-health emergency could harm tourism in the nearby Manú National Park and Tambopata National Reserve.

Because of the growing concerns over mercury exposure, Fernandez is leading a project at Wake Forest University in Winston-Salem, North Carolina, to study the metal's effects on human and environmental health in the Amazon. As director of the Center for Amazonian Scientific Innovation at Wake Forest, Fernandez will lead a team of US researchers who are collaborating with colleagues at the Peruvian Amazon Research Institute and the National Amazonian University of Madre de Dios. ■

DAN COLLINS/AFP/GETTY



## HUMAN GENETICS

# Ambitious plan to synthesize human genome unveiled

*Some admire project to make genome from scratch; others say it hasn't justified its aims.*

BY EWEN CALLAWAY

Proposals for a large public-private initiative to synthesize an entire human genome from scratch — an effort that could take a decade and require billions of dollars in technological development — were formally unveiled on 2 June, almost a month after they were first aired at a secretive meeting.

Proponents of the effort, named 'Human Genome Project—Write' (HGP-write), write in *Science* that US\$100 million from a range of funding sources would help to get their vision off the ground<sup>1</sup>. The team is led by synthetic biologist Jef Boeke at New York University; genome scientist George Church at Harvard Medical School in Boston, Massachusetts; and Andrew Hessel, a futurist at the commercial design studio Autodesk Research in San Rafael, California.

But the idea — which essentially aims to develop technologies that reduce the cost of DNA synthesis — has not met with universal excitement among researchers. To some, the proposal to create a human genome is praiseworthy for its ambition and sheerchutzpah: at present, only tiny bacterial genomes and a portion of a yeast genome have been made from scratch. But others feel that HGP-write represents a needless centralization of work that is already taking place in companies trying to lower the price of synthesizing strings of DNA (see 'Falling costs'). Some of HGP-write's proponents have financial stakes in those firms, which include Gen9 in Cambridge, Massachusetts.

"My first thought was 'so what,'" says Martin Fussenegger, a synthetic biologist at the Swiss Federal Institute of Technology in Zurich. "I personally think this will happen naturally. It's just a matter of price at the end."

Others think that the project should be delayed until its leaders can win broader support for the idea. In an e-mail sent to reporters, synthetic biologist Drew Endy, at Stanford University in California, and religion scholar Laurie Zoloth, at Northwestern University in Evanston, Illinois, say that the HGP-write team has not properly justified its aims, and that the project should be abandoned. "We are still waiting for a serious public debate with participation from a broad range of people," they say.

Endy and Zoloth had already questioned the scientific rationale for synthesizing a human

genome in May, when HGP-write was first aired at an invitation-only meeting at Harvard University that was attended by more than 100 scientists, entrepreneurs, lawyers and ethicists. The closed nature of the meeting also attracted

College London. "If you want to do this, it's going to be on the same scale as the Human Genome Project, it's going to need some big funding agencies and hundreds and hundreds of researchers around the world."

Ellis and others worry that a centralized project that explicitly focuses only on the human genome might needlessly narrow the products of the effort. But Boeke — who in 2014 reported synthesizing a 270,000-base-pair yeast chromosome<sup>2</sup> — brushes that objection aside. He envisions HGP-write eventually producing synthetic genomes from mice, microbes and all sorts of organisms.

"Tangible products may be slow to follow at first, but writing DNA more cheaply and at large scale will make researchers more efficient and comprehensive in their work, leading to practically unlimited potential for indirect products," adds Danielle Tullman-Ercek, a biochemical engineer at the University of California, Berkeley.

Cheaper DNA synthesis isn't the only thing that stands in the way of writing a human genome that could function inside a cell. Ellis says that there are no methods for inserting very large pieces of DNA into a mammalian cell and making them function normally, and researchers have little clue how to design a complex genome that has anything more than trivial changes to an existing one.

Freemont worries that commercial bodies such as DNA-synthesis companies may stake a claim on the project. "I think it's good if this is an open, publicly funded initiative," he says.

Boeke would prefer that there be no intellectual-property restrictions on the products of HGP-write, as is the case with his synthetic yeast genome project. But, he says, the "chances are good" that companies involved in HGP-write will be granted such rights "to get the job done."

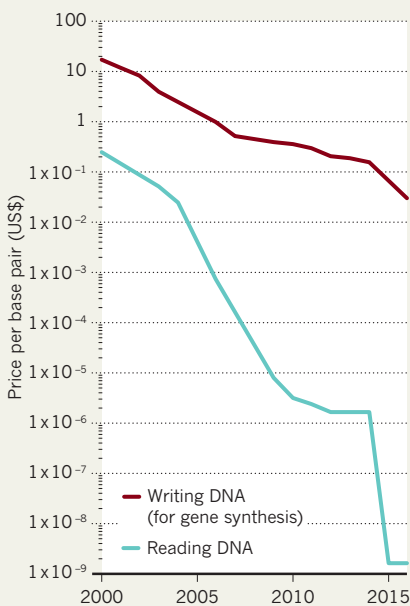
In the *Science* article, the HGP-write team says that it will seek broader public buy-in before beginning work. Boeke says that much of the Harvard meeting centred on the ethics of the project. And he says that the synthetic cells will be engineered to make reproduction impossible. "We're not trying to make an army of clones or start a new era of eugenics. That is not the plan." ■

1. Boeke, J. D. et al. *Science* <http://dx.doi.org/10.1126/science.aaf6850> (2016).

2. Annaluru, N. et al. *Science* **344**, 55–58 (2014).

## FALLING COSTS

The price of 'writing' DNA is falling more slowly than the price of reading it.



criticism: Church told the health and medicine news service Stat that this was because the paper describing the effort was under embargo. "There was a lot of confusion on the day about what was going on," says Tom Ellis, a synthetic biologist at Imperial College London.

The three-page announcement of HGP-write fills in some detail. It notes that current technologies are both too expensive and too primitive to synthesize the 3-billion-base-pair human genome. The team calls for a series of pilot projects, including synthesizing much shorter segments of the genome and making slimmed-down chromosomes to do specific tasks, to make its eventual goal doable. The whole project should require less than \$3 billion (the price of the publicly funded Human Genome Project), the researchers say.

"I think it's a brilliant project," says Paul Freemont, a structural biologist at Imperial

## POLICY

## UK loses track of its policy research

*Report slams government for missing studies.*

BY DANIEL CRESSEY

The UK government doesn't know how much policy-linked research it has commissioned, or how much of such research has been published. That is the stark conclusion of an independent inquiry, published on 2 June, which details confusion about the status of research produced for government departments in areas ranging from social policy to climate change.

The inquiry was carried out by Stephen Sedley, a judge, law professor and trustee of Sense About Science, the London-based science-advocacy group that published the report. He spoke to government advisers, civil servants and researchers, and used multiple freedom-of-information requests to find out how much research commissioned by the government gets published.

According to official estimates, the government spends around £2.5 billion (US\$3.6 billion) a year commissioning research linked to policy issues. But, Sedley says, it has "no comprehensive account" of how much is commissioned or published.

Just 4 out of 24 government departments told Sedley that they kept a centralized database of commissioned research. Others could not provide a list of the studies that they carried out or commissioned. Many departments said that it would be too costly to provide the information, because it was held in many different files and locations.

Civil servants told Sedley that they often waste time trying to find past studies. The report also notes several cases in which the publication of reports has been delayed owing to "political concerns about the implications of the research" — including work on drug policy and immigration.

"The fact that a few departments do maintain a research register, handle awkward findings and publish promptly exposes the excuses of those that don't," said Tracey Brown, director of Sense About Science, in a statement. The report calls for a central register of all government-commissioned research, a commitment to prompt publication, and routine publication of any work that has been used to inform government policy. ■



The late rock-art specialist Mike Morwood at Liang Bua cave, where his team discovered *Homo floresiensis*.

## PALAEOANTHROPOLOGY

## Hobbit relatives hint at family tree

*Possible ancestors of *Homo floresiensis* found after long hunt.*

BY EWEN CALLAWAY

More than a decade after the discovery that a diminutive relative of modern humans once lived on the Indonesian island of Flores, Gerrit van den Bergh was losing faith that he would find any clues to the ancestors of the 'hobbit'. It was October 2014, and for four years he had co-led an industrial-scale excavation near the cave where the metre-tall skeleton had been found. Then, weeks before packing it in for the year, a local worker found a 700,000-year-old molar. More teeth and a partial jaw quickly followed.

"We had given up hope we would find anything, then it was 'bingo!'," says van den Bergh, a palaeontologist at the University of Wollongong, Australia, whose team reports the finds in two papers in this issue (G.D. van den Bergh *et al. Nature* 534, 245–248; 2016; and A. Brumm *et al. Nature* 534, 249–253; 2016). "We had this enormous party. We had a cow slaughter and there was dancing. It was marvellous."

The unusually petite jaw and teeth are from

at least one adult and two children — the first possible ancestors of *Homo floresiensis* ever to be discovered — and resemble the hobbit remains found on the island, which are between 60,000 and 100,000 years old.

The jaw and teeth address two questions that have dogged the study of the species — where did it come from and how did it get so small? But as with all things hobbit, there is little consensus among researchers, who say that firm conclusions require more fossils.

The hobbit's discovery in 2003 in Liang Bua cave, by a team led by the late Australia-based rock-art specialist Mike Morwood, was an instant sensation. But its place in the human family tree is contentious. Morwood's team proposed that it was a shrunken *Homo erectus*, the same species that probably evolved into *Homo sapiens* in Africa and that roamed as far as Europe and Asia. Other scientists who have examined features of *H. floresiensis*, such as its long, flat feet, think that it descended from a smaller, more primitive human relative such as *Homo habilis* or even *Australopithecus*,

ACHMAD IBRAHIM/AP



known only from remains in sub-Saharan Africa.

Seeking the hobbit's ancestors, in 2004, Morwood's team returned to a site 74 kilometres from Liang Bua called Mata Menge, where elephant bones and tools had been found in the 1960s. The dig started small, but in 2010 the team scaled up. Bulldozers cleared an area of 2,000 metres square, and more than 100 locals then dug for 6 days a week using chisels and hammers. They found hundreds of stone tools, thousands of fossils from animals such as crocodiles, rats and komodo dragons, but no hominin bones.

By then ill with advanced prostate cancer, Morwood visited the area for the last time in 2012. "He really made an effort to walk through the site, you could see he was in pain, but he was so detailed-minded," van den Bergh says. "He increased the pressure to dig more holes and go faster. He really wanted to find them."

Morwood, who died in 2013 before the teeth and jawbone were found, is an author on the *Nature* papers, which were co-led by scientists based in Japan, Australia and Indonesia.

The team concludes that the jaw excavated at Mata Menge is from an adult (its wisdom tooth had erupted) who was even smaller than the hobbit, and that two canines are the milk teeth of two different children. The thin jaw looks more like that of *H. erectus* and *H. floresiensis*

than the beefier jaws of more primitive hominins such as *H. habilis*. The square-shaped teeth are intermediate between *H. erectus* and *H. floresiensis*.

**"We had given up hope we would find anything, then it was 'bingo!'"**

One tooth and the rock around it led the team to estimate that the remains are some 700,000 years old. The oldest artefacts in the region, meanwhile, suggest that a group of *Homo erectus* arrived on Flores about one million years ago, says van den Bergh.

#### DWARFED BY DIET

He and his team note that the remains point to large-bodied *H. erectus* as the likeliest ancestor of the hobbit, and propose that it shrank in just a few hundred thousand years to cope with the meagre resources on Flores. Elephants and other large creatures have been known to shrink over time to cope with the lack of food typical of islands, and red deer on the island of Jersey in the English Channel shrank to one-sixth of their original size in just 6,000 years, says van den Bergh.

Both Fred Spoor, a palaeontologist at University College London, and palaeoanthropologist Chris Stringer at London's Natural History Museum agree that *H. erectus* is now the best fit for the hobbit's ancestor, although

Stringer isn't so sure that the shrinkage happened on Flores. It's just as likely that the hobbit emerged on another island, such as Sulawesi, and then moved to Flores, he says.

But William Jungers, a palaeoanthropologist at Stony Brook University in New York, says that the fossils are not complete enough to favour the *H. erectus* origin: "I don't believe these scrappy new dental specimens inform the competing hypotheses for the origin of the species one way or another."

A small river that leads down a hill deposited the sandstone in which the teeth and jaw were found, and van den Bergh expects that more hominin remains lie there. His colleagues, meanwhile, have found stone tools in Sulawesi, north of Flores. For once, the prospect of more hobbits isn't looking so bleak. ■ [SEE EDITORIAL P.151](#) [AND NEWS & VIEWS P.188](#)

#### CORRECTION

In the News Feature 'South Korea's Nobel dream' (*Nature* **534**, 20–23; 2016), one paragraph incorrectly gave numbers in billions instead of trillions of won. In fact, 63.7 trillion won was spent on R&D, 49.2 trillion of which came from private enterprise; and 11.2 trillion won was spent on basic research.



# Delhi's

BY MEERA SUBRAMANIAN

## *India's capital scrambles to tackle its epic pollution problems.*

**O**n winter nights, New Delhi burns with innumerable fires. Flames flicker along pavements and street corners, where the destitute huddle to stay warm and cook their suppers, while night watchmen stand guard next to their own small blazes outside private homes. The rising plumes of smoke mingle with exhaust and dust stirred up by overloaded trucks that rumble down roads blanketed in fog. The mixture melds into a nearly opaque substance that leaves a metallic taste on the tongue. Overhead, there is not a single star to be seen.

With dawn comes a hint of warmth, but the sunlight remains hidden by haze. A hopelessly optimistic sign — “Make Delhi Pollution-Free” — is lashed to a metal cage that protects a young sapling, its withered leaves caked with dust.

The grime is the most obvious part of the pollution that plagues India's capital region and its 25 million people. Less discernible are the airborne particles smaller than 2.5 micrometres in diameter, known as  $PM_{2.5}$  — the most harmful size range. Just a fraction of the diameter of a human hair and astoundingly aerodynamic,  $PM_{2.5}$  can penetrate deep into the body, reaching the recesses of the lungs. The particles are a nasty amalgam of pollutants both natural and increasingly anthropogenic, generated from sources within the city's boundaries and hundreds of kilometres away. The World Health Organization (WHO) declares that no amount of this pollutant is safe to breathe.

Two years ago, Delhi had the highest  $PM_{2.5}$  levels of 1,600 cities surveyed by the WHO. Last month, in an updated and expanded inventory<sup>1</sup>, Delhi retained its status as the most polluted of the world's largest cities, with an annual  $PM_{2.5}$  average of 122 micrograms per cubic metre ( $\mu g m^{-3}$ ) — three times the permitted Indian standard and greatly exceeding the WHO standard of  $10 \mu g m^{-3}$ . The pollution, which comes mainly from combustion of wood, coal, gas, diesel and

**A brown cloud of pollutants hangs over the outskirts of New Delhi.**





# deadly air

crop residue, is worst in the winter, when wood-burning peaks and cold-weather inversions trap pollutants close to the ground and cause spikes in the daily average of above  $600 \mu\text{g m}^{-3}$ . Late last year, the levels prompted the Delhi High Court to declare the city a “gas chamber”.

The observed  $\text{PM}_{2.5}$  amounts are estimated to cause as many as 16,000 premature deaths and 6 million asthma attacks<sup>2</sup> in Delhi annually, shaving around 6 years off the life expectancy of city residents<sup>3</sup>. Although the WHO data have brought attention to Delhi, the problem is global: according to the agency, particulate pollution affects more people than any other pollutant on Earth.

## AIR PATROL

Delhi, like Beijing, Mexico City, London and Los Angeles, is struggling to reduce pollution, even as its population swells. Researchers and the government are trying to construct a detailed breakdown of the different pollution sources, and authorities are experimenting with ways to mitigate the damage, from restricting when people can drive to shutting down power plants. But India faces unique challenges. Its population is concentrated in the north, an area geographically prone to pollution, and its people have aspirations for development. There is a growing middle class hungry to own cars, and one-fifth of the population merely wants access to basic electricity. Those facts threaten to compromise Delhi's efforts to improve environmental quality.

“The reality is that the pollution in Delhi is very complex. There are a lot of sources. It varies from season. It varies by time of day. It varies by neighbourhood,” says Namit Arora, a member of the pollution task force of the Delhi Dialogue Commission, a government initiative in the

city. But he insists that the city can make progress. “We can act, and we need to act, on multiple fronts simultaneously.”

Delhi is trying to do just that. Long before it was saddled with the mantle of having some of the world's worst air, the city and the Supreme Court of India took several steps to alleviate pollution. Vehicle emissions came down in the early 2000s, thanks to decisions to remove lead from petrol, improve vehicle-emission standards and pull old commercial vehicles off the roads. Around the same time, the city implemented a monumental conversion of the public-transportation fleet, including buses and the city's zippy three-wheel auto-rickshaws, away from gasoline and diesel engines to ones fuelled by cleaner compressed natural gas. In 2002, the Delhi Metro subway system opened its first line, improving public-transportation options. All but two of the coal-based thermal power plants in the city were converted to natural gas, and many industries, including brick kilns, were moved beyond Delhi's bounds.

These efforts reaped big gains, yet they have been offset by the incessant growth of the megacity. Since 2000, Delhi's population has nearly doubled. And the number of vehicles has almost tripled, from 3 million to more than 9 million, according to the government.

Determining what creates hazardous  $\text{PM}_{2.5}$  is a crucial step in reducing it, yet past studies have varied markedly because they have used different methods and relied on limited data. Filling some of the gaps is a report released in January by the Indian government and the Indian Institute of Technology (IIT) Kanpur<sup>4</sup>, which took a more comprehensive approach to investigating the causes of Delhi's poor air quality. Some sources contribute throughout the year, such as pollution from vehicles, diesel generators, construction dust, biomass and



coal combustion and industries. Others are seasonal: dry summer dust blowing in from nearby deserts, autumn crop burning and Diwali holiday fireworks, and the warming fires that make the city glow come winter (see 'Poison stew').

Vehicle emissions are constant, and with all those belching tailpipes in sight every day, they often capture the attention of pollution-fighting officials. In January, Delhi implemented an odd-even programme that allowed car owners to drive only every other day, as dictated by their vehicle's number plate. When the 15-day experiment came to a close, a few hundred Delhiites took to the streets to praise the initiative and rally for continued efforts. Families, musicians, passionate teenagers, costumed 20-somethings and activists all gathered near Jantar Mantar, a cluster of monuments built in the 1700s to study skies that were then crowded with visible stars. But over the temporary stage set up for the event, an air monitor showed that  $\text{PM}_{2.5}$  levels were hovering around  $184 \mu\text{g m}^{-3}$ , a level that warrants staying indoors to reduce exposure.

Although levels remained well above acceptable during the odd-even experiment, several researchers declared it a success in both lowering emissions and, perhaps more importantly, raising public awareness. "People are willing to display the civic sense we thought didn't exist here," says Arora, who was at the demonstration wearing a shirt emblazoned with the words "Help Delhi Breathe" and an image that was half flower bloom, half gas mask.

Still, government officials and researchers admit that this approach has limited long-term potential because of how difficult it is to enforce the ban. During January's trial, people were already talking about buying a second car as a workaround. After the odd-even effort was repeated in April, the government estimated that there were half a million more vehicles on the roads than during January's trial, according to local media, which suggests that people were skirting the rules.

The odd-even policy gets a lot of attention, "but that is not the solution," says Ashwani Kumar, who was secretary of the environment of the Delhi Pollution Control Committee (DPCC) at the time. "A typical democratic society cannot depend on a 'don'ts' approach," he says. Making it more expensive and inconvenient to drive a car, for example, would naturally spur people to use public transportation. "It has to be based on incentive and disincentive; otherwise, it's easy to find loopholes and descend into a quagmire of corruption."

## GOING PUBLIC

Although parts of Delhi's public-transportation system are impressive, others are lacking. The Delhi Metro is an efficient, extensive electric-rail system with more than 200 kilometres of lines, and it continues to expand. But the Indian media has reported that the government's plans to increase the number of buses in Delhi have been plagued by delays, and that a pilot programme for dedicated bus lanes was met with so much public resistance that the lanes are now being dismantled.

"If the public-transportation system is robust, and it's made safe and comfortable and reliable, people will automatically switch," says Sarath Guttikunda, director of the independent research group UrbanEmissions.info, which is registered in Delhi.

Nudging people onto buses and subways will deal with only part of the transportation pollution problem. Although vehicles as a group contribute up to about one-quarter of  $\text{PM}_{2.5}$  in Delhi<sup>5</sup>, the fraction generated by heavy-duty freight trucks is twice that of cars. To ease the impact of the estimated tens of thousands of trucks that move through the city daily, the Supreme Court has implemented new taxes on them, and Delhi is adding bypass highways.

The other crucial ingredient is the type of fuel that goes into vehicles. Given that diesel engines produce much more particulate matter than ones that run on petrol, the rising percentage of luxury diesel cars is a troubling trend. To try to stem the sales, the government temporarily banned registration of diesel vehicles with larger engines earlier this year, according to Indian media.

Delhi's emissions standards for vehicles are more stringent than those in the rest of India, but they still lag far behind those in Europe. And

the discrepancy between city and national standards means that many vehicles operating in Delhi spew emissions from lower standard vehicles and fuel obtained beyond city limits.

Yet all this attention on vehicles is somewhat misplaced because they are not the biggest source of particulate pollution. "If the goal is to reduce PM, we need to go beyond traffic," says atmospheric scientist Pallavi Pant of the University of Massachusetts–Amherst, who studies Delhi's air quality.

Just a few steps away from the busy roads, on a broken stretch of pavement, a woman tends a cooking pot perched on three stones, a wood fire burning below. Across South Asia, more than one-quarter of the outdoor air pollution comes from these traditional stoves<sup>6</sup>. In urban



Ever-present construction and smoke-belching vehicles are major contributors to air pollution.

## As it seeks cleaner air, Delhi has to confront the desperate desire for development.

CHRIS STOWERS/PANOS

Delhi, only about one in ten households still relies on smoky stoves that use wood, dung or kerosene, but they still contribute a substantial part of the city's  $\text{PM}_{2.5}$  burden. One study found<sup>7</sup> that the emissions from fires used domestically for cooking and warmth rival those from the electricity sector, from brick kilns and from industry.

Delhi's government and the country's Supreme Court have tried to limit some of these sources, but there is one major factor that they cannot control: the city's location, far from the cleansing breezes of the ocean. From the west, smoke from agricultural crop burning and dust storms from the Thar Desert blow into the city. And from the north, cold fronts sweep down from the Himalayas, locking pollutants in with winter weather inversions. When it comes to air pollution, the city is "geographically disadvantaged", says M. P. George, an environmental scientist with the DPCC.

Although Delhi can't fend off desert dust storms, it can control construction dust generated by its never-ending building spree. Officials are attempting to decrease air pollution from construction by improving and enforcing rules, such as requirements to cover building sites and trucks to stop dust from blowing away.

The problems that Delhi faces also afflict other cities in the region, creating a vast brown cloud that in satellite images seems to smother much of South Asia. About one-third of Delhi's particulate pollution comes from sources outside the city, according to the IIT-Kanpur study<sup>4</sup>.

"It's not just a Delhi problem; it's a regional problem," says Milind Kandlikar, who researches development and environmental issues at the University of British Columbia in Vancouver, Canada. India's northern



cities have come to dominate the WHO's list of most-polluted cities, with half of the top 20 all located in the region. Delhi's efforts will falter unless the rest of the country also steps up, says Kandlikar. "Everybody else has to be involved. This is not going to go away easily."

But political conflicts are threatening the chances for broad-scale pollution-control efforts. There is currently a fractious tension between the national government, run by Prime Minister Narendra Modi's Bharatiya Janata Party, and Delhi's ruling Aam Aadmi Party — a situation that hinders efforts to develop a unified strategy to deal with air pollution. In lieu of any cooperation, the Supreme Court continues to have an integral role in improving air quality by directing the govern-

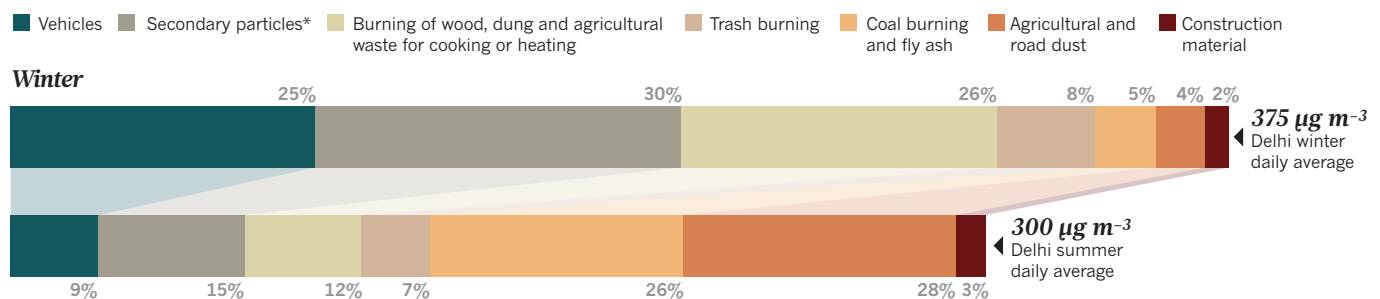
ment to restrict emissions from sources such as vehicles and power plants. "My reading is that the politicians are often grateful to the courts for acting, because many of these measures are unpopular," says Arora. With so many sources of air pollution, city officials would ideally like to know which programmes will work best — so the Delhi Dialogue Commission task force recruited IBM Research-India to develop computer models to forecast pollution levels.

odd-even schemes, and the Delhi government has teamed up with the University of Chicago's Delhi-based academic centre to launch a design competition called the Urban Labs Innovation Challenge: Delhi, which aims to crowdsource ideas for improving air and water quality. In March, it received hundreds of submissions, including ideas for promoting rooftop solar panels and creating viable alternatives to burning waste and crops. Prize money of up to US\$300,000 will fund design pilots.

In Delhi and around the world, citizens, governments and researchers are all demanding more air-quality data, which indicates an interest in knowing the enemy. Information from government monitors is publicly available, but the interfaces are often clunky. This has given rise

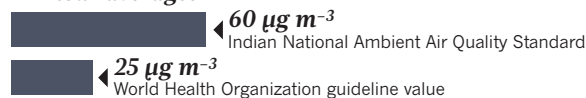
## Poison stew

Delhi has the highest particulate air-pollution readings of any megacity. A study released this year by the Indian Institute of Technology Kanpur<sup>4</sup> found that different sources dominate in winter and summer for particles smaller than 2.5 micrometres, known as PM<sub>2.5</sub>.



\*Formed mostly from sulfur dioxide and NO<sub>x</sub> produced principally from vehicles, industry and power generation.

### 24-hour averages:



ment to restrict emissions from sources such as vehicles and power plants. "My reading is that the politicians are often grateful to the courts for acting, because many of these measures are unpopular," says Arora.

With so many sources of air pollution, city officials would ideally like to know which programmes will work best — so the Delhi Dialogue Commission task force recruited IBM Research-India to develop computer models to forecast pollution levels.

"The idea is to create a simulation framework which will allow them to evaluate different policies and see their impact before they actually implement them," says Ashish Verma, a senior manager with IBM's Smarter Planet team in Delhi. Although the modelling is still in progress, early results reinforce how important the weather is. This suggests that air-pollution controls might be most effective if they take advantage of specific weather conditions, for example, by closing power plants or limiting traffic during the worst pollution-trapping weather inversions.

### CITIZEN ACTION

As it seeks cleaner air, Delhi has to confront the country's desperate desire for development, which includes providing electricity to the roughly 240 million people who still lack it. India pledged to expand its renewable energy capacity aggressively, as part of the national plan that it submitted during the United Nations climate-treaty negotiations last year. But the plan also defends India's right to use fossil fuels, stating that "in order to secure reliable, adequate and affordable supply of electricity, coal will continue to dominate power generation in future".

Both the climate plan and government officials such as Kumar insist that India need not follow conventional modes of development that rely on fossil fuels. "There is no way we can adopt the technology trajectory of the developed countries who continue to be the biggest polluters in terms of per capita," Kumar says. But it is not clear whether India will be able to leap-frog past the most polluting forms of energy.

One encouraging sign in Delhi is that air quality is now part of an active city-wide conversation. Full-page newspaper ads solicit citizen input on

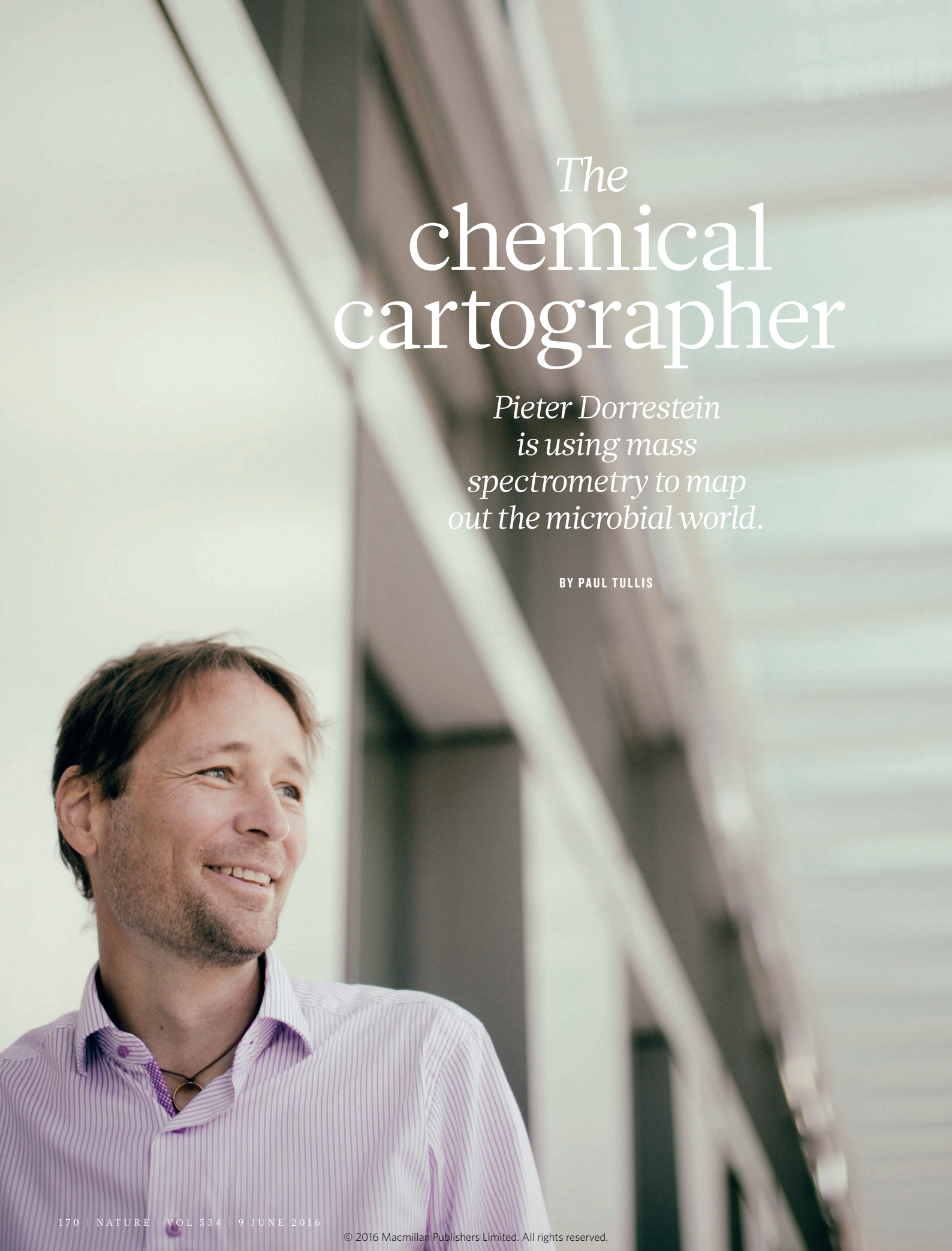
to new alternatives such as IndiaSpend, an independent media outlet that also conducts 'sensor journalism', in which readings from its own pollution monitors are made available in a user-friendly format. The sheer increase in global monitoring — the WHO's database doubled in two years because many more cities had begun to monitor their air — offers an opportunity for more-comprehensive studies going forward, including in regions that have previously been neglected.

Meanwhile, concerns about pollution are hard to escape. In upscale Delhi markets, vendors hawk air masks and purifiers, and parents can purchase nebulizers decorated with cute animals to appeal to children with asthma. Doctors have been known to advise patients with lung ailments to leave the city. "Picking up a life is not so easy," pleads one mother. "What are we supposed to do?"

Many Delhiites tend to swing between despair and hope — just as the skies go through their own cycles. By late March, the weather starts to shift and the winds pick up. It's as if the city's windows are thrown open, allowing fresh breezes to blow through. July brings the monsoon rains and washes much of the danger out of the sky, down towards the Yamuna River, which carries some of the burden away. For a few months, citizens can step into the night, tilt their heads back and once again enjoy the stars overhead. ■

**Meera Subramanian** is a journalist in Cape Cod, Massachusetts, and is the author of *A River Runs Again: India's Natural World in Crisis*, from the Barren Cliffs of Rajasthan to the Farmlands of Karnataka.

1. World Health Organization. *WHO's Urban Ambient Air Pollution Database — Update 2016* (WHO, 2016).
2. Guttikunda, S. K. & Goel, R. *Environ. Dev.* **6**, 8–20 (2013).
3. Ghude, S. D. *et al. Geophys. Res. Lett.* **43**, 4650–4658 (2016).
4. Sharma, M. & Dikshit, O. *Comprehensive Study on Air Pollution and Green House Gases (GHGs) in Delhi* (Indian Institute of Technology Kanpur, 2016).
5. Goel, R. & Pant, P. *Econ. Polit. Wkly* **51**, 41–45 (2016).
6. Smith, K. R. *et al. Annu. Rev. Public Health* **35**, 185–206 (2014).
7. Guttikunda, S. K. & Calori, G. *Atmos. Environ.* **67**, 101–111 (2013).



# *The* chemical cartographer

*Pieter Dorrestein  
is using mass  
spectrometry to map  
out the microbial world.*

BY PAUL TULLIS



A part from the treadmill desk, Pieter Dorrestein's office at the University of California, San Diego (UCSD), is unremarkable: there is a circular table with chairs around it, bookshelves lined with journals, papers and books, and a couple of plaques honouring him and his work.

But Dorrestein likes to offer visitors a closer look. On his computer screen, he pulls up a 3D rendering of the space. Four figures seated around the table — one of whom is Dorrestein — look as if they've been splashed with brightly coloured paint. To produce the image, researchers swabbed every surface in the room, including the people, several hundred times, then analysed the swabs with mass spectrometry to identify the chemicals present.

The picture reveals a lot about the space, and the people in it. Two of Dorrestein's co-workers are heavy coffee drinkers: caffeine is splashed across their hands and faces (as well as on a sizeable spot on the floor — a remnant of an old spill). Dorrestein does not drink coffee, but has left traces of himself everywhere, from personal-care products to a common sweetener that he wasn't even aware he'd consumed. He was also surprised to find the insect repellent DEET on many of the surfaces that he had touched; he hadn't used the chemical in at least six months.

Then there were signatures of the office's other inhabitants: the microbes that reside on human skin. Dorrestein has been using mass spectrometry to look at the small molecules, or metabolites, produced by these microbes, and to get a clearer picture of how microorganisms form communities and interact — with other microbes, with their human hosts and with the environments that they all inhabit.

He has analysed microbial communities from plants, seawater, remote tribes, diseased human lungs and more, in an effort to listen in on their chemical conversations: how they tell one another of good or bad places to colonize, or fight over territory. The work could identify previously unknown microbes and useful molecules that they make, such as antibiotics.

"The applications are broad," says Katie Pollard, a comparative genomicist at the Gladstone Institutes at the University of California, San Francisco. Because many microbes cannot be cultured and studied directly, she explains, "these approaches that assay them *in situ* are totally game-changing". They also directly address some of the main goals outlined in the US\$521-million National Microbiome Initiative, announced by the White House's Office of Science and Technology Policy last month. Dorrestein was present for the announcement.

In this fast-moving field, Dorrestein has set himself apart by building useful tools and productive collaborations. "Pieter is genuinely interested and very creative," says Janet Jansson, division director of biological sciences at Pacific Northwest National Laboratory in Richland, Washington. In April, she visited UCSD, and Dorrestein asked whether he could swab her hand for one of his studies. "I said, 'Oh! I want to do that! I want to be involved in that study!'" Jansson recalls. "It's interesting and exciting science that people want to participate in."

## ROCK AND ROLL

Dorrestein grew up in the Netherlands, and became obsessed with rock-climbing when he visited family friends in Tucson, Arizona, at the age of 16. Faced with the flatness of his homeland, he applied to Northern Arizona University in Flagstaff, in large part because of its proximity to the many stone towers of the Four Corners region, where Arizona meets New Mexico, Colorado and Utah. He studied geology and chemistry, but intended to pursue his passion for climbing. Shortly after graduating in 1998, however, an experience on the 900-metre-tall face of El Capitan in Yosemite, California, made him think again.

He was clinging to the rock about 50 metres above his last anchoring point, and realized that if he were to lose his grip, he would drop

100 metres before his safety line tautened and slammed him into the granite. It wasn't fear, he says, but rather his lack of it that troubled him. "I thought, if I keep doing this, it won't be a good ending," he recalls. "So I rappelled down."

He drove home to Flagstaff that day, and started filling out applications to graduate school. He ended up at Cornell University in Ithaca, New York, studying how microbes produce small molecules such as vitamin B1. It was here that he was first introduced to mass spectrometry.

Mass spectrometry generally involves breaking complex molecules apart, ionizing them and measuring the mass of the resulting fragments, which can be used to calculate the composition of the starting molecules. Dorrestein uses the analogy of a bar code — mass spectrometry creates a unique identifier for each chemical in a sample.

Spurred by his interest in the technology, he went on to do a post-doc in the lab of Neil Kelleher, a chemical biologist at the University of Illinois at Urbana-Champaign. Kelleher was pioneering efforts to do 'top-down' mass spectrometry, in which intact, rather than digested, proteins are put directly into the mass spec. The approach allows researchers to identify small modifications made to proteins, but the process is slow. Within two months of his arrival in Illinois, Dorrestein had developed a speedier approach that allowed him to examine certain large enzymes systematically<sup>1</sup>. "We boiled down years of work into days, basically," Dorrestein says. He ended up co-authoring 17 papers in 2 years. "Pieter has that unusual combination of creativity and drive, along with an incredible ability to finish projects," says Kelleher, who is now at Northwestern University in Evanston, Illinois.

Dorrestein joined the faculty at UCSD in 2006 — but things really kicked off for him when Palmer Taylor, then dean of the university's school of pharmacology, authorized the purchase of a MALDI-TOF mass spectrometer (matrix-assisted laser desorption/ionization time of flight), which would allow Dorrestein to do mass-spectrometry imaging. "That changed the whole world around," he says.

## SPACE CRUSADERS

As well as identifying molecules in a sample, mass-spectrometry imaging provides spatial information. MALDI-TOF uses a laser to heat up and ionize molecules. By scanning that laser across a 2D sample, researchers can capture an 'image' that shows exactly where different molecules in the sample reside. The technique can be used to identify and locate biomarkers in slices of tumours, but with his interest in microbes, Dorrestein wondered whether he could take colonies of bacteria on a Petri dish and scan them directly to see the metabolites they produce.

No one had ever tried it. Dorrestein suspects that they were afraid of getting their expensive mass spectrometers dirty — "and this is as dirty as it comes, putting microbes directly into the instrument". So he tried a simple experiment, asking an undergraduate student, Sara Weitz, to scan a colony of *Bacillus* bacteria.

The images generated "weren't the prettiest", Dorrestein says, but they indicated that the process worked. He sent them to Paul Straight, a microbiologist who had just joined the faculty at Texas A&M University in College Station. "I'm pretty sure his jaw dropped," Dorrestein says. Together, the two teams used mass-spectrometry imaging on colonies of *Bacillus subtilis* and *Streptomyces coelicolor* grown next to one another. By exploring the spaces where the colonies interacted, they were able to identify molecules that the microbes use to compete with each other<sup>2</sup>.

Actually visualizing this microbial arms race, Dorrestein says, makes him think back to 1928, when Alexander Fleming isolated penicillin from a mould that was killing bacteria on a dish. Mass-spectrometry imaging could quickly reveal the chemistries of such interactions, and perhaps speed up the search for new antibiotics.

Dorrestein decided to shift his lab to focus almost exclusively on these methods. He was still an early-career investigator, and almost everybody he knew discouraged him from taking such a big risk. But Taylor pushed him to apply for tenure right away. "Pieter's potential to think outside the box in the analytical and computational arenas was immediately evident," Taylor says. "His research took off very rapidly."

The problem with looking at dirty samples is that they produce messy data. Scanning microbial landscapes produces thousands of bar

Pieter Dorrestein's methods could reveal what microbes do in complex communities.

SANDY HUFFAKER

codes, but it's largely unknown what they correspond to; they haven't been annotated. "It's the equivalent of looking under the lamp post," Dorrestein says: one can only 'see' the molecules that have been identified before, and the vast majority haven't. This is currently a big challenge for the field, says Jansson. "It's possible to analyse features by mass spec, but still very difficult to identify what those features are."

To help to make sense of the heaps of data, Dorrestein worked with Nuno Bandeira, a computational biologist at UCSD, on an approach that classifies bar codes and the molecules to which they correspond according to their relationships with other annotated molecules<sup>3</sup>. This allows researchers to start predicting, computationally, the structures and functions of thousands of metabolites. But there's still a dearth of annotation: although thousands of people worldwide conduct mass-spectrometry research, most annotate only the few molecules that they're interested in.

So, beginning in 2014, Dorrestein and graduate student Mingxun Wang from Bandeira's lab started to develop a way to crowdsource annotation. They launched the Global Natural Product Social Molecular Networking website, a repository and data-analysis tool that enables researchers to uncover relationships between related molecules, group similar ones together and compare data sets. "This is something he's brought to the field that has really helped," says Jansson.

## TEAM WORK

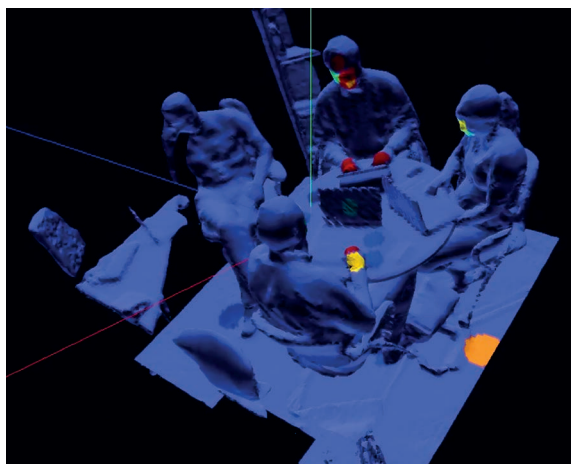
One of the keys to Dorrestein's success has been his collaborations. Rob Knight, a leader in microbiome DNA and RNA sequencing, works just across the quad from Dorrestein's office. They've teamed up to blend sequencing with mass spectrometry. Last year, a postdoc in Dorrestein's lab, Amina Bouslimani, took swabs from one male and one female volunteer, at 400 spots on their bodies — twice. One swab from each spot went to Knight's lab so that the microbes in it could be sequenced, and the other went for mass spectrometry to identify the chemicals, natural and artificial, that coexist with the microorganisms.

The participants had refrained from showering or using cosmetics for three days, but the chemical signatures from the hundreds of different types of microbe in the samples were overwhelmed by chemicals from beauty and hygiene products<sup>4</sup>. Still, the researchers did find correlations between microbe communities and local chemistries: for example, the bacteria found in the vaginal area were correlated with molecules associated with inflammation. Such connections, Dorrestein says, could be used to generate hypotheses about host–microbe interactions.

Bouslimani is now analysing samples from volunteers' hands and from personal items such as their mobile phones. The work, which has not yet been published, has shown that people leave persistent chemical signatures on the objects that they touch — like those in the image of Dorrestein's office.

Bouslimani and Dorrestein think that this could have applications in forensic science. A suspect could be swabbed to determine whether the chemical signature of his or her skin matches that at a crime scene. Or in the absence of DNA or fingerprint evidence, the chemicals that a criminal leaves behind could help to provide a lifestyle profile: a composite sketch of the products that they use and the mixture of microbes they carry. "Maybe the chemical signature can help the investigator narrow down who was there," says Bouslimani.

Last year, Dorrestein teamed up with microbiologist Maria Dominguez-Bello of New York University and several others who wanted to see what human skin and its microbial diversity look like when people grow up free of the trappings of the developed world. They



Caffeine appears as coloured spots in a 3D visualization.

collected samples from some remote tribes — one near Manaus, Brazil, and Tanzania's Hadza people — and compared them with swabs from non-tribal people near the collection sites. Using Dorrestein's mass-spectrometry techniques, they've found that people in the tribes have more-diverse microbial communities and skin chemistry than those living a more modern lifestyle. The ongoing work is serving up some surprises too, says Dorrestein. People from one village in Brazil had a range of pharmaceuticals on their skin, indicating that they had more contact with outsiders than previously suspected.

Dorrestein has a way of leaning forward and almost standing on his toes in excitement when he talks about

the technology and how it might help to assess the health of oceans, or improve efficiency in agriculture, a major contributor to greenhouse-gas emissions. But when asked how he chooses projects to pursue, it's work on human health that he mentions first. "To us, that's a really obvious, direct application of this — we want to help patients," he says.

Dorrestein teamed up with Knight, Doug Conrad — director of UCSD's adult cystic fibrosis clinic — and others to develop a rapid microbial diagnostic test. Cystic fibrosis causes a build-up of mucus in the lungs, which can periodically become infected with bacteria. These infections require aggressive treatment with antibiotics — and sometimes the bacteria can develop resistance. Dorrestein and his collaborators have shown<sup>5</sup> how analysing mass-spectrometry data on a phlegm sample from someone with cystic fibrosis can identify microbial communities that standard medical culturing techniques miss.

Louis-Félix Nothias-Scaglia, a postdoc who joined Dorrestein's lab this year, is mapping the skin of people with psoriasis, a condition thought to be triggered by an overactive immune system. If molecules produced by certain bacteria are present when the condition flares up but not when the skin is healthy, Nothias-Scaglia explains, they might point to drugs that could treat or even prevent the disease. Even being able to use microbial changes to predict when a flare-up is coming would enable patients to reduce their use of immune-suppressing drugs.

Turning such data-intensive techniques into standard lab tests will be a challenge. "Cynics would say it's too complicated, it's never gonna go anywhere," says Conrad. "To a certain extent, I can understand that. But that's a good way to keep going the way things are."

Dorrestein definitely wants to change the way things are, particularly for the blossoming field of microbiome research. He views the discipline as passing through phases: the first has centred on determining the identity of microbes. The second phase is working out what they're doing, using techniques such as mass spectrometry.

What drives the establishment of these communities? What metabolic processes are under way, and how do they interact with each other and with a host? "If you fundamentally understand that," Dorrestein says, "you can start to take control of it." And that's the third phase, he says — taking control. By monitoring microbial communities, is it possible to add the necessary ingredients to change a person's health, their mood, their athletic performance? Dorrestein thinks that the answers to these questions are right in front of him. He just has to look a little closer. ■

Paul Tullis is a freelance journalist in Los Angeles, California.

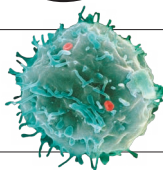
1. Dorrestein, P. C. *et al. Biochemistry* **45**, 12756–12766 (2006).
2. Yang, Y.-L., Xu, Y., Straight, P. & Dorrestein, P. C. *Nature Chem. Biol.* **5**, 885–887 (2009).
3. Watrous, J. *et al. Proc. Natl Acad. Sci. USA* **109**, E1743–E1752 (2012).
4. Bouslimani, A. *et al. Proc. Natl Acad. Sci. USA* **112**, E2120–E2129 (2015).
5. Quinn, R. A. *et al. ISME J.* **10**, 1483–1498 (2016).

DATA: A. BOUSLIMANI/C. KAPONO/T. ALEXANDROV; VIDEO FRAMEGRAB: P. DORRESTEIN/UCSD/COLLABORATIVE MASS SPECTROMETRY INNOVATION CENTER



# COMMENT

**STEM CELLS** Analysis defines four versions of 'stemness' **p.176**



**MILITARY SCIENCE** A tour of conflict's wilder shores **p.178**

**CONSERVATION** A call to audit the impacts of ivory burning on poaching **p.179**

**OBITUARY** Ilkka Hanski, ecologist of fragmented habitats, remembered **p.180**

ILLUSTRATION BY DAVID PARKINS



## Lessons from researcher rehab

Common compliance situations can get good researchers into trouble, warn **James M. DuBois** and colleagues.

In 2013, we launched a training programme that no scientist wants to list on their CV. Participants are referred to it by their home institutions, usually after having their research privileges suspended. Several times a year, a small group of researchers arrives in St Louis, Missouri, for a course that we designed to help participants regain their research privileges. We are grant-funded faculty members with backgrounds in psychology and research ethics. The programme — initially called RePAIR, now called the

Professionalism and Integrity Program, or PI Program — was developed with funding from the US National Institutes of Health, and participants pay a fee to attend.

About half of our participants are enrolled after a failure of oversight that resulted in the publication of false data or faulty consent. Other common reasons include plagiarism and documentation problems. Researchers are often referred on multiple grounds (see 'Why attendees enrolled'; and Supplementary Information; [go.nature.com/25ao1sd](http://go.nature.com/25ao1sd)).

Many arrive convinced that they have been misjudged. One said, "The programme sounds like it's for criminals — not me."

Three intense days later, attitudes change. Participants generally express gratitude for the programme. A year later, follow-up surveys indicate that the vast majority have changed how they work. We, the programme instructors, have also adjusted our research practices: we now appreciate how easy it is to run afoul of rules when busy schedules collide with complex projects. ▶

► Before arriving at our small group workshop, participants complete a battery of assessments. On the final day, they write a professional-development plan, which outlines strategies to, say, hold regular team meetings, seek further training on compliance rules, or restructure workloads. Over the next three months, we conduct an average of three coaching telephone calls with alumni as they implement their plans.

We began our programme not knowing who would attend, or even whether what we were doing was a good idea. Three years in, we're convinced that it is worthwhile. We have now trained 39 researchers from 24 institutions. Researchers in our programme do not display personality traits that are distinct from the general population of scientists. We believe that most researchers may be susceptible, and that the busiest ones are most likely to err.

### WHAT WENT WRONG?

There are high-profile cases of serial fraudsters who have consciously built their careers on fabricated data and who, some research suggests, have personality disorders<sup>1</sup>. We do not encounter such individuals in our programme. In general, we work with talented faculty members who seek to do good research and whom institutions wish to retain.

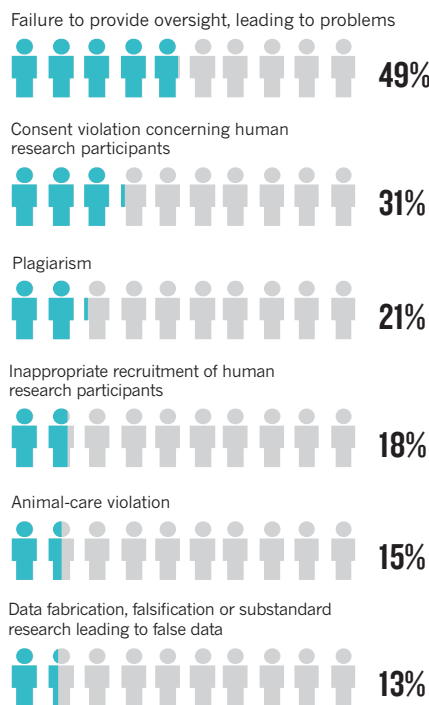
To understand what led to these individuals' mistakes, we draw on several data sources. Over the past two years, we have surveyed participants on factors such as their knowledge of research rules, job-related stress, how much an individual uses cognitive distortion to justify compliance breaches<sup>2</sup>, and their use of strategies that improve professional decision-making<sup>3</sup>. We compared their survey scores with those of a national sample of 400 researchers funded by the US National Institutes of Health. Our participants did not differ from this sample's mean scores on any measures thought to be connected to compliance or research integrity.

We also profiled participants' work-related attributes. The Clifton StrengthsFinder is a test to identify employees' strongest talents<sup>4</sup>. It is widely used by organizational psychologists in corporate settings, and its validity and reliability has been documented in diverse professions. Most of our participants had top scores in 2 of 34 possible areas: achiever and learner. This is not surprising, given that all participants had doctoral degrees and had been in full-time employment in research institutions. However, test results showed a dearth of compliance-related talents, such as focus, discipline and consistency.

**"Our programme participants were highly successful researchers."**

### WHY ATTENDEES ENROLLED

Frequent reasons behind researchers' referrals to the Professionalism and Integrity Program (many are referred for more than one reason).



Although we do not have a national comparison sample for this test, a meta-analysis suggests that most scientists would share such a profile. Among the most common traits associated with scientific creativity are: being less conventional, being more open to new experiences, and ambitiousness<sup>5</sup>.

During day two of the workshop, participants share with the group what led to their referral. Assurances of confidentiality and the presence of others in similar situations mean that participants are markedly forthright. We reflected on their stories and other information gathered to contemplate why they found themselves in trouble. For most, we identified multiple causes (see 'Why researchers stumbled').

Three causes played a part in most cases: paying too little attention to details or oversight; being unsure about relevant rules; and not prioritizing compliance. All these could be attributed to other, more basic causes. For example, many participants provided too little oversight of their teams because they were overextended or understaffed. People sometimes were unsure of rules after moving into a new area of research. They also encountered regulations that had grown more complex since they completed their training.

### THREE MYTHS ABOUT MISCONDUCT

Our experience with programme participants challenges several misconceptions about misconduct.

**Only bad apples get into trouble.** We do not want to minimize the seriousness of participants' violations. Institutions face large fines for non-compliance with federal rules; deficient research hurts the quality of scientific literature, and human and animal subjects may be put at risk. Nevertheless, participants' infractions rarely resulted from a conscious intent to mislead or break rules.

Often, programme participants had in their research tried to ensure that the fundamental purpose of compliance was satisfied — for example, that research participants understood and freely consented to study conditions, that animals were kept pain free, or that data supported conclusions. They did so, however, in ways that fell short of full compliance. As one participant put it: "Prior to this situation, I tried to follow the spirit of the law — now I try to follow the letter of the law."

Though our participants were not different from other researchers in terms of their knowledge of rules, knowledge deficits sometimes played a big part in their situations. Most researchers function fairly well within complex regulatory systems, but problems arose when they entered new territory — for example, running their first clinical trial involving drugs or devices subject to regulations of the Food and Drug Administration.

The role of culture in research compliance is rarely discussed. For example, in the United States, plagiarism is taken much more seriously than violation of rules about how credit for work is assigned, but such common values are rarely expressed explicitly. Research-ethics programmes often attach the same significance to all rules, and may leave some scientists ignorant of unspoken norms.

A slight majority of our participants came to the United States from elsewhere; roughly double the expected proportion on the basis of the demographics of the US scientific workforce. (Our previous studies indicate that nation of birth is a stronger predictor of scores on key measures than is nation of training.) Although many participants were accustomed to US compliance requirements, they sometimes employed or mentored people who were not. Some of these junior lab members made poor assumptions about appropriate processes that participants failed to correct. For example, one participant did not review a postdoc's data and analyses because he felt doing so would imply mistrust. We suggested that the participant frame his reviews of lab members' work as modelling appropriate behaviour, explicitly telling postdocs that they would also be expected to provide this kind of oversight if they headed a lab.

**Scientific skills are enough to do good science.** The idea behind the StrengthsFinder assessment is that people tend to be strong in some areas but not in others. Scientists have long had to master skills associated with, for



instance, artists (creativity) and accountants (attention to detail). In an increasingly competitive environment, requirements for both talents have increased, and scientists must also have the good communication and negotiation skills often associated with politicians. Building teams with complementary skills is ideal, but it is not always easy or even possible when resources are limited.

How researchers communicate with others (team members and institutional officials) matters greatly. Communication should be clear, balanced and non-threatening, especially when an institution questions a researcher's compliance. For example, informing institutional officials that your grants pay their salaries is rarely effective in resolving concerns.

When situations demand skills that a team does not have, compensating strategies may be helpful<sup>3</sup>. For example, standard operating procedures (SOPs) or checklists can facilitate the consistent performance of compliance tasks — such as getting approval for each animal-research protocol — even when tasks or team members are new, or when people are tired. The principal investigator who creates SOPs for compliance and data integrity also sends the message that these are as important as the science itself. Even with

SOPs, principal investigators must actively oversee compliance activities and ensure that staff are adequately trained, competent and diligent. Regular meetings can help. Principal investigators are accountable for the integrity of their research projects — they must find ways to assure it even when they themselves are not detail-oriented.

**The more publications and grants the better.** By the metrics that institutions use to reward success, our programme participants were highly successful researchers; they had received many grants and published many papers. Yet, becoming overextended was a common reason why they failed to adequately oversee research. It may also have led them to make compliance a low priority. People who are too busy must triage, and what scientist wants to prioritize checking patient signatures above data gathering?

Principal investigators should protect themselves and their labs by taking on no more projects than they can responsibly oversee and adequately staff. In general, to do more, researchers need more resources — space, trained staff and mentoring when moving into new areas. A clear lesson from our programme is that compliance requires individual integrity as well as departmental

and institutional support. Scientists become overextended in part because their institutions value large numbers of projects.

## GOOD INVESTMENT

Before its launch, the programme was criticized on principle: why allocate resources to assist rule breakers to regain their research privileges?

We think these are resources well spent. Questionable research practices are much more widespread than we would like to believe<sup>6</sup>. Following the workshop, our participants demonstrate more positive attitudes toward compliance, improved problem-solving skills and better lab-management habits. Proactive, preventive training is often recommended as a way to boost research integrity, but there is little evidence that increasing one-size-fits-all training changes behaviour. Certainly, sending someone back to repeat a standard course after a finding of misconduct seems unlikely to help<sup>7</sup>.

In our view, intense, individualized training following a breach can be remarkably effective. And it is unquestionably much more cost efficient than letting problems fester until even bigger problems arise for investigators and institutions<sup>8</sup>. Our participants have gone from analysing their own lapses to customizing solutions, such as holding more face-to-face meetings or developing SOPs.









Our experience with the course has made us more compassionate to the participants, and more cautious about our own behaviours. The message that we want to send is this: unless you are careful, it could happen to you. Learning to do compliance well ensures data integrity and protects human and animal subjects; it also protects the careers of researchers and their colleagues. ■

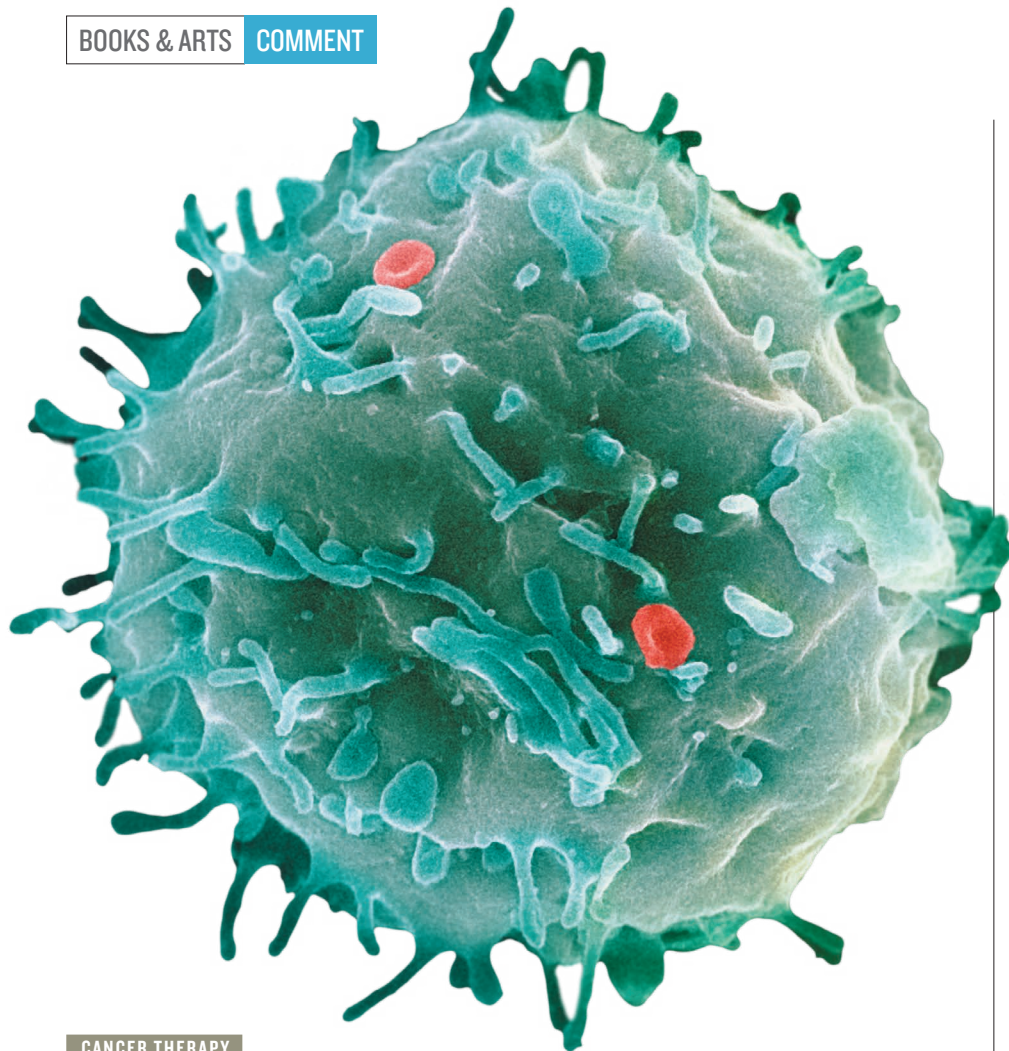
**James M. DuBois** is director of the Center for Clinical & Research Ethics at Washington University School of Medicine, St Louis, Missouri, USA. **John T. Chibnall** is professor of psychiatry, **Raymond Tait** is professor of psychiatry and vice-president of research, and **Jillon Vander Wal** is professor of psychology, at Saint Louis University, St Louis, Missouri, USA.  
e-mail: [duboisjm@wustl.edu](mailto:duboisjm@wustl.edu)

- Kornfeld, D. S. *Acad. Med.* **87**, 877–882 (2012).
- DuBois, J. M., Chibnall, J. T. & Gibbs, J. *Sci. Eng. Ethics* <http://dx.doi.org/10.1007/s11948-015-9681-x> (2015).
- DuBois, J. M. *et al. Sci. Eng. Ethics* **22**, 391–416 (2015).
- Asplund, J., Lopez, S. J., Hodges, T. & Harter, J. *The Clifton StrengthsFinder 2.0 Technical Report: Development and Validation* (The Gallup Organization, 2009).
- Feist, G. J. *Personal. Social Psychol. Rev.* **2**, 290–309 (1998).
- Martinson, B. C., Anderson, M. S. & de Vries, R. *Nature* **435**, 737–738 (2005).
- Antes, A. L. *et al. Acad. Med.* **85**, 519–526 (2010).
- Michalek, A. M., Hutson, A. D., Wicher, C. P. & Trump, D. L. *PLoS Med.* **8**, e1000318 (2010).

## WHY RESEARCHERS STUMBLER

Instructors on the Professionalism and Integrity Program assessed underlying causes (often more than one) for researchers' lapses.

Proximate cause	Ultimate cause of researcher lapse	% of participants
Lack of attention	Overextended, not detail-oriented or distracted by personal problems.	 72%
Unsure of rules	An increase in regulations since researcher began career, lack of mentoring or cultural differences.	 56%
Did not prioritize compliance	Failed to recognize seriousness of violations, biased thinking or cultural differences.	 56%
Relationship problems, political tensions	Communicated aggressively or worked with difficult personalities.	 36%
Staff lacked adequate training or integrity	Failed to provide adequate training, did not create culture of compliance in lab or had difficulty hiring individuals.	 28%
Poor communication	Failed to hold regular meetings with research team.	 26%
Ambition	Driven personality, desire for promotion or competition for funding.	 21%
Conflicting roles (physician–scientist)	Interacted with individuals as both patients and research participants.	 21%
Did not anticipate consequences	Failed to consider ways a project could go wrong.	 13%
Lack of resources	Inadequate institutional investment in researcher's programme.	 10%
Followed poor instructions	Rigid hierarchy in research programme and the absence of positive mentors to consult.	 10%



A stem cell seen under a scanning electron microscope.

its title, builds a much broader framework for understanding the

biology of stem cells of all types.

Central to CSC theory is the observation that not all tumour cells are equal. The bulk of a tumour consists of short-lived proliferative cells and differentiated cells. But some tumour cells seem to be the malignant equivalents of tissue stem cells. Much as normal stem cells maintain healthy organs by producing new tissue cells, CSCs drive the persistence of malignant tumours by producing new cancer cells.

CSC theory tacitly assumes that CSCs carry the armaments of normal stem cells: they are built to last a lifetime; are resilient to many kinds of chemical or physical insult; and can 'slumber' for prolonged periods. CSCs would thus be capable of surviving chemotherapy and radiation, explaining why local recurrence is the almost inevitable outcome of such treatment. And metastases that sometimes appear many years after the excision of a primary tumour would be caused by quiescent CSCs that have wandered to distant sites. Thus, CSC theory explains why cancer patients can never be considered cured, even when a treatment outcome seems encouraging. Most important, CSC theory promises the development of innovative treatments, aimed not at reducing the bulk of a tumour, but at taking out its 'beating heart': the cancer stem cells.

Laplane starts her comprehensive overview with a description of how the popularity of CSC theory has exploded over the past two decades, driven by the rapid development in cell-sorting technology. She then gives an insightful historical account starting from nineteenth-century giants including Theodor Schwann and Rudolf Virchow. Almost as an aside, she describes the work of Leroy Stevens and Barry Pierce on transplantation of cells from teratocarcinomas, eerie growths that can contain any type of tissue, including teeth and hair. This eventually led to the discovery and 'domestication' of their healthy physiological counterparts, the embryonic stem cells that generate all tissue types in early embryos.

She discusses, too, how Canadian stem-cell biologist John Dick revived the field in the 1990s by developing ways to study the behaviour of different types of human leukaemia cell by transplanting them into mice. He merged this with modern insights into the nature of the healthy stem cells that give rise to all types of blood cell (haematopoietic stem cells).

He discovered that within a patient, a minority of leukaemia cells closely resemble healthy stem cells, but

➔ **NATURE.COM**

For more on science in culture see: [nature.com/booksandarts](http://nature.com/booksandarts)

## CANCER THERAPY

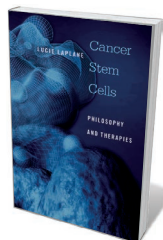
# Defining stemness

Hans Clevers admires an analysis of stem-cell science that sharpens up some of the fuzziness in the field.

I have always felt uncomfortable about the concepts and definitions that we use in the stem-cell field. Some of the arguments seem circular; observation and assumption are not well separated. I once asked a colleague for their best definition of a stem cell. The answer: a cell that can self-renew. What, then, is self-renewal? The immediate reply: what stem cells do.

Fuzziness in stem-cell concepts and definitions has significant consequences. It affects how we design, conduct and interpret experiments, how we communicate our discoveries and, ultimately, how we design therapies aimed at supporting the regenerative capacity of healthy stem cells or eradicating those that fuel the growth of tumours. Despite these concerns, as an experimentalist I could never put my finger on where exactly scientific common sense is failing.

Enter Lucie Laplane and her book



**Cancer Stem Cells: Philosophy and Therapies**  
LUCIE LAPLANE  
Harvard University Press: 2016.

literature — historical and scientific — she interviewed some of the leading international stem-cell researchers and clinicians. She discussed her emerging insights with fellow philosophers and science historians. Starting from an interest in cancer stem cells (CSCs), the book, despite

### *Cancer Stem Cells.*

Trained as a science philosopher, Laplane also spent time at the bench in two stem-cell labs. Her book is the culmination of a six-year effort to describe and structure the philosophical underpinnings of stem-cell science. In addition to absorbing essentially all the relevant experimental



rather than producing normal blood cells, they continuously produce leukaemia cells. Thus, Laplane concludes that the definition and study of CSCs is inseparable from that of normal stem cells.

Laplane defines a stem cell as one “capable of self-renewal and of differentiation”, where self-renewal is the ability to recreate one copy of itself on division. Laplane reveals how for some time the field was describing two very different entities as CSCs: the cells from which tumours originate; and those inside tumours that drive their growth in the long term. Apples and oranges. The field has also wrestled with the gold-standard assay introduced by Dick: the transplantation of sorted human cancer cells into mice, where the cells that grow out into tumours are considered CSCs. It is

**“Fuzziness in stem-cell concepts has significant consequences, but I could never put my finger on where scientific common sense is failing.”**

arguable what this ‘surrogate’ assay actually measures. Does the outgrowth of heavily manipulated cancer cells in mice really reflect the behaviour of the same cells in the original tumour?

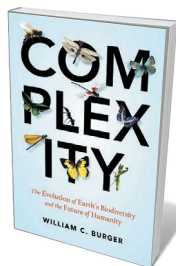
Laplane acknowledges these issues. In a second edition, she may want to touch on how researchers use genetic marking of stem cells to trace their derivatives through solid tissues.

Laplane’s rigorous analyses unveil deep semantic and conceptual problems in the field. She arrives at a framework of four possible versions of ‘stemness’: two intrinsic, two extrinsic. She suggests that it can be categorical (an intrinsic property of a stem cell, independent of its environment); dispositional (an intrinsic property of a stem cell that emerges only in the right environment); relational (an extrinsic property induced in a cell that would otherwise be a non-stem cell by its microenvironment); or systemic (an extrinsic property of a system such as tissue, rather than an individual cell).

I suspect that there is no current consensus on where to fit even the best-studied stem-cell types into this framework. Yet Laplane’s stemness framework should be of great value. It will help to clarify definitions and concepts, even if it only provides solid ground from which to disagree. Moreover, the framework can readily be applied to experimentation. A philosopher may indeed have straightened out the stem-cell field. ■

**Hans Clevers** investigates stem cells and cancer at the Hubrecht Institute in Utrecht, the Netherlands. He is also the research director of the national paediatric oncology hospital Princess Máxima Center in Utrecht. e-mail: h.clevers@hubrecht.eu

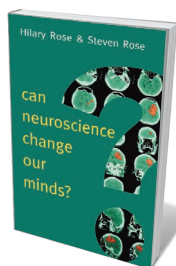
## Books in brief



### Complexity: The Evolution of Earth's Biodiversity and the Future of Humanity

William C. Burger PROMETHEUS (2016)

Botanist William Burger conducts a grand tour of life’s complexity, emphasizing cooperation and symbiosis in evolutionary history. He segues deftly from the towering success of beetles and bacteria to the formation of new species and the distribution of biodiversity. The story culminates with humanity’s cognitive and cultural hegemony. But however ascendant we are as a species, Burger dispassionately notes, our explosive global population growth and overuse of resources mirror the behaviour of locust swarms.



### Can Neuroscience Change Our Minds?

Hilary Rose and Steven Rose POLITY (2016)

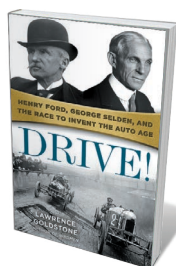
The seepage of neuroscience into economics and policy should be deeply questioned, argue sociologist of science Hilary Rose and neuroscientist Steven Rose in this crisp, astringent analysis. In a historically and scientifically contextualized critique of this “data-rich and theory-poor” discipline, they examine claims made for the US and European ‘big brain’ projects, and for the findings that feed into UK policy on child-rearing and early education. Ultimately, they aver, neuroscience can indeed change our minds — but social and political understanding of the issues must be factored in.



### What a Fish Knows: The Inner Lives of Our Underwater Cousins

Jonathan Balcombe FARRAR, STRAUS AND GIROUX (2016)

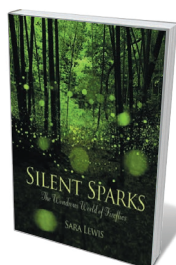
More than 30,000 species of fish — about half of all vertebrates — roam global waters. And as ethologist Jonathan Balcombe notes in this engrossing study, breakthroughs are revealing sophisticated piscine behaviours. Balcombe glides from perception and cognition to tool use, pausing at marvels such as ocular migration in flounders and the capacity of the frillfin goby (*Bathygobius soporator*) to memorize the topography of the intertidal zone. Yet, he argues, the over-exploitation of wild stocks, notably of apex predators such as tuna, points to the need for change on moral as well as ecological grounds.



### Drive! Henry Ford, George Selden, and the Race to Invent the Auto Age

Lawrence Goldstone BALLANTINE (2016)

Historian Lawrence Goldstone follows the momentous patent war that ended in 1911, when George Selden’s case for a patent on a “road carriage” powered by internal combustion was broken by arch-industrialist Henry Ford, who adapted existing technology to craft the wildly successful Model T. Goldstone weaves in accounts of European innovators such as Karl Benz, and road races such as the 1907 Peking-to-Paris dash. But as the market-savvy maverick who “did not so much create demand as anticipate it”, Ford dominates the story.



### Silent Sparks: The Wondrous World of Fireflies

Sara Lewis PRINCETON UNIVERSITY PRESS (2016)

The pulsing glow of massed fireflies is a nocturnal wonder of nature. Biologist Sara Lewis has spent decades studying these beetles of the family Lampyridae, which spans nearly 2,000 species. Here she expounds on firefly metamorphosis, courtship, reproduction and bioluminescence — from the exquisite anatomy of the *Photinus* firefly’s lantern to the chemical ‘light switch’ that enables flash control. (A field guide to North American fireflies is included.) An illuminating peek into a fascinating corner of field biology. **Barbara Kiser**



Soldiers benefit from — and are subject to — huge amounts of research.

#### MILITARY SCIENCE

# Among the warriors

Sharon Weinberger finds much to amuse and disturb in Mary Roach's tour of conflict's wilder shores.

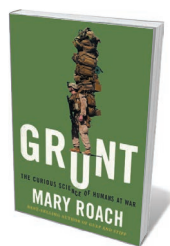
In *Grunt*, Mary Roach reveals herself as the kind of writer you want at a nap-inducing military press conference. In a culture where dead civilians are 'collateral damage' and air strikes are 'kinetic operations', Roach has a way of peeling back the euphemisms to get to some of the true horrors of war. Unbelievably, she often manages to make it funny.

The book, a tour of the scientific world of "humans at war", sees Roach uncovering the Medicare reimbursement code for maggots and learning how combat medics practise treating an evisceration (the stand-in for faeces-filled intestines involves dyed oatmeal and worse). To find humour in the carnage of war, a writer must walk a fine line, and most of the time Roach does so deftly. Her previous works published by W. W. Norton have focused on sex (*Bonk*, 2008), space travel (*Packing for Mars*, 2010), cadavers (*Stiff*, 2003) and other subjects ripe for snappy one-liners. War is more challenging. Yet even her chapter on genital injuries — urotrauma — strikes an appropriate balance. It is at once harrowing, fascinating and depressing.

Roach is at her best capturing the rational absurdities of the US military, an institution at perpetual war. For instance, there is a task-group dedicated to the "hook-and-loop fastener" (preferred by snipers over the noisier Velcro, which can reveal their position). If the

devil is in the details, then war is all devils: "Only a military clothing designer's portfolio would include a mitten that accommodates a lone forefinger in firing position."

Roach's writing has been criticized as superficial, but that is not its greatest weakness here. Rather, there is something disturbing about approaching military science as if it were all so awesome. A public-affairs officer is "likeable", a medical researcher "gorgeous". Who knew that everyone in and around the military was so "droll and adorable"? Roach may never have seen a military representative stonewalling a legitimate inquiry, as happened when the military faced allegations of neglect at the Walter Reed Army Medical Center in Washington DC in the 2000s. Or



**Grunt: The Curious Science of Humans at War**  
MARY ROACH  
W. W. Norton: 2016.

perhaps she is hesitant to admit it, because that is not her style. More worryingly, Roach's happy world of military science does not include cases in which the military treated people as human guinea pigs, such as in radiation experiments at the height of the cold war. Roach does give

a glimpse of the military's underbelly. Her chapter on the implications of diarrhoea as an impediment to combat readiness takes her to the secretive Camp Lemonnier in Djibouti. This passage says much about the US military, which allowed a writer interested in loose stools onto this drone-operations hub, but has been reluctant to let in national-security reporters. She corners one of the base's mysterious "special operators"; he reveals little about operations, but speaks freely on his bowel movements.

There are some significant gaps. Roach does not, she concedes, cover post-traumatic stress disorder (PTSD), "not because PTSD doesn't deserve coverage, but because it has so much, and so much of it is so very good". That is a shame: Roach's unique voice might have done great things for the mouse models that military researchers use to study psychological trauma (imagine a very timid mouse being menaced by a large one bred for aggression). It's also troubling because PTSD, along with traumatic brain injury, is an important part of current military science. Among other subjects, the military is looking at the links between traumatic brain injuries and dementia (see *Nature* 477, 390–393; 2011).

Also missing are the central questions that make military science so fascinating. For example, how can the exigencies of national security advance knowledge in areas such as trauma medicine, yet create decades of controversy in others, such as nonlethal weapons? Descriptions of wounds inflicted by roadside bombs in Iraq and Afghanistan are *Grunt's* strongest passages, but the later chapters on the science of humans under water and the perils and benefits of flies on the battlefield jump around. It's hard to escape vertigo as we skip from penis transplants to shark repellent. Roach misses an opportunity to examine how military science has morphed along with war.

The most telling part of *Grunt* comes early in a chapter on heat, when Roach notes that "genetic differences in thermoregulation" are important "given our seemingly permanent posture of fighting extremism in the Middle East". The medical science of war has never been static, but a constant reflection of where and how the military fights. Here, Roach has inadvertently gone to the heart of the perspective that is missing from *Grunt*. In the era of traumatic brain injury, lost limbs and PTSD, the focus of scientists studying humans at war has evolved from getting soldiers to survive battle to getting them to survive peace. ■

Sharon Weinberger is a fellow at the Radcliffe Institute for Advanced Study at Harvard University in Cambridge, Massachusetts, and author of a forthcoming book on the US Defense Advanced Research Projects Agency. e-mail: sharonweinberger@gmail.com



# Correspondence

## Track the impact of Kenya's ivory burn

Kenya's government delivered a powerful message against elephant poaching and the illegal ivory trade on 30 April by burning 105 tonnes of ivory, worth up to US\$220 million. With stockpile destruction on the rise, it is important to evaluate the impact of this strategy on elephant populations.

Since 1989, 21 countries have burned or crushed 263 tonnes of ivory — most of it (86%) in the past 5 years (see [go.nature.com/ivory](http://go.nature.com/ivory)). However, there is no published evidence so far that these events reduce poaching.

Destroying ivory stockpiles risks a perverse outcome: ivory becomes rarer, fetching higher prices and increasing poaching and illegal stockpiling (see M. 't Sas-Rolfes *et al.* *Pachyderm* **55**, 62–77; 2014). This has prompted calls by some for a highly controlled legal ivory trade to secure elephant populations (J. F. Walker and D. Stiles *Science* **328**, 1633–1634; 2010) — an option that ivory destruction removes.

It is therefore crucial to track the effects of Kenya's largest-ever ivory burn. Time is short and the stakes are high.

**Duan Biggs\*** *University of Queensland, Brisbane, Australia.*  
[ancientantwren@gmail.com](mailto:ancientantwren@gmail.com)  
\*On behalf of 4 correspondents (see [go.nature.com/1rt2mhe](http://go.nature.com/1rt2mhe) for full list).

## China's primates: EU can't have it all ways

We are concerned about the prospect of China becoming a world leader in research involving non-human primates, given the country's comparatively weak regulatory system and ethical framework (see *Nature* **532**, 281; 2016 and *Nature* **532**, 300–302; 2016).

China's relative freedom from the “ethical pressure” you mention makes it attractive to researchers working on primates.

But animal studies that could fail the harm–benefit evaluation in many Western regulatory systems should not be allowed — or actively encouraged — to take place elsewhere. Far from putting researchers under negative ethical pressure, the project-authorization process in the European Union was set up with full input from scientists and is often held up (by them) as an appropriate safeguard to promote good quality, ethically conducted science and good animal welfare (*Nature* **521**, 7; 2015). You cannot have it both ways.

Rather than exploiting weaker animal-research regulations, we argue that more effort should be invested in developing and validating alternative technologies to avoid or reduce the use of non-human primates.  
**Penny Hawkins, Paul Littlefair**  
*Royal Society for the Prevention of Cruelty to Animals, Southwater, UK.*  
[penny.hawkins@rspca.org.uk](mailto:penny.hawkins@rspca.org.uk)

## China's primates: preserve wild species

China is being put forward as a world leader in primate biomedical research (see *Nature* **532**, 300–302; 2016), even while its wild populations of primates are being lost at an alarming rate because of illegal activities and poor conservation practice.

Take rhesus macaques (*Macaca mulatta*), the species most frequently used in biomedical research. The wild population of these primates in China was estimated at about 200,000 in 2008, with a further 40,000 kept in breeding centres (see [go.nature.com/1sl2bqx](http://go.nature.com/1sl2bqx)). This number of captive animals has since significantly increased and may, along with the 20,000 exported each year, include individuals that were bred outside captivity (see X. Hao *Cell* **129**, 1033–1036; 2007).

Despite the changes that China is making on paper to improve its conservation policies, the

declining state of its 19 native primate species conveys a different story. These animals are disappearing because of habitat disturbance, illegal export and hunting — including for traditional medicine.

The country seems to us to be more concerned with increasing its reputation in biomedical primate research. That reputation will be boosted by the large input of government funding and by Western researchers flocking in for the reasons you mention.

China's position on conservation issues and on primate welfare should not be skated over. Animals are not exploitable, and wild populations should not be an afterthought.  
**Alison M. Behie, Colin P. Groves**  
*The Australian National University, Canberra, Australia.*  
[alison.behie@anu.edu.au](mailto:alison.behie@anu.edu.au)

## Supervise Chinese environment policy

China's latest five-year plan shifts its environmental law away from a pollution-control system and towards one that manages environmental quality (see *Nature* **531**, 524–525; 2016). Regional efforts will now be subject to greater oversight to ensure that improvements are implemented across the country and to prevent local corruption.

Under the plan, provincial environmental-protection departments will be responsible for unifying local monitoring and inspection programmes and for eliminating protectionism in local governments (see B. Zhang and C. Cao *Nature* **517**, 433–434; 2015). China's Ministry of Environmental Protection has already established separate environmental-management departments for water, air and soil.

In my view, strict national supervision would help to keep these regional reforms on track and to make them more effective. The US Environmental Protection Agency, for example,

manages pollution through permits and defined standards. Ten regional offices work with individual states to implement these regulations. The agency can revoke state programmes that fail to fulfil their responsibilities.

In China, the unified supervision of local monitoring and inspection by provincial environmental-protection departments, which then report to the ministry, is an important step towards improving environmental quality. However, each of the links in this chain must strictly enforce the regulations and work with the rest to clean up the country's environment.

**Bo Zhang** *Information Center, Ministry of Environmental Protection, China.*  
[zhangbo@mep.gov.cn](mailto:zhangbo@mep.gov.cn)  
Disclaimer declared (see [go.nature.com/24m3usa](http://go.nature.com/24m3usa) for details).

## Change of identity is not in the air

Change is indeed in the air for many butterflies — at least in their ecology, if not in their outer appearance (see ‘Change is in the air’ *Nature* **532**, 403–404; 2016). However, the example you picture is not the European species *Hesperia comma*, but the North American *Epargyreus clarus* — both of which have the same vernacular name of silver-spotted skipper.

**Josef Settele** *Helmholtz Centre for Environmental Research — UFZ, Halle, Germany.*  
[josef.settele@ufz.de](mailto:josef.settele@ufz.de)

### CORRECTION

The Correspondence by P. Dobosz and J. Zawila-Niedzwiecki (*Nature* **532**, 441; 2016) incorrectly described some Polish universities as “engaging” anti-scientific speakers. In fact, the speakers either hired university premises or participated in discussions at university conferences.

# Ilkka Hanski

## (1953–2016)

Population ecologist who modelled how species cope with habitat loss.

Ecologist Ilkka Hanski's pioneering work changed our understanding of how biodiversity is maintained. Combining mathematical modelling and long-term data from the wild, he developed metapopulation theory. This predicts the degree of habitat fragmentation beyond which a species will go extinct.

Hanski's 1999 book *Metapopulation Ecology* (Oxford University Press) became a cornerstone for researchers in population biology, conservation biology and landscape ecology. He identified the genetic basis of traits that underpin survival in fragmented habitats. Most recently, he demonstrated with colleagues that an increasing prevalence of inflammatory diseases is associated with declining biodiversity.

Hanski, who died on 10 May, was born in 1953 in Lempäälä, Finland. As a child he collected butterflies at his grandparents' house in southeastern Finland. Decades later, Hanski reflected that many of his most successful research projects were inspired by these outdoor adventures and by the encouragement that he received from Esko Suomalainen, a geneticist at the University of Helsinki whom he had contacted after finding a rare butterfly at his grandparents' house.

Hanski studied biology at the University of Helsinki and gained his doctorate, on the community ecology of dung beetles, from the University of Oxford in 1979. He found that most dung-beetle species clump together, with particular species common in some parts but scarce or absent in others.

Before the 1970s, ecologists paid little attention to whether populations were distributed continuously or in many local patches or both. Throughout his career, Hanski was intrigued by the ecology and evolution of species found in islands — naturally fragmented habitats. In 1969, population biologist Richard Levins introduced the concept of a metapopulation — a 'population of populations' — species living in networks of habitat patches such as cowpats or islands, work that Hanski built on.

Returning to Helsinki after his doctorate, Hanski continued to develop models for metapopulation survival. By the late 1980s, he was ready to test his predictions in the field, but which insect to study, and where? He was inspired during a fortuitous visit by renowned population biologist Paul Ehrlich of Stanford University in California. Discussions about Ehrlich's research on Edith's checkerspot butterfly (*Euphydryas editha*) prompted Hanski



to choose the Glanville fritillary butterfly (*Melitaea cinxia*) in the Åland Islands off southern Finland to test his predictions.

In the early 1990s, he set out to map all suitable habitat patches for *M. cinxia* in Åland. His effort has grown into a database of more than 4,000 localities. These places are checked each year for *M. cinxia* and its larval host plants — as well as the parasitoids and pathogens of each.

This has since become one of the most important model systems in population biology. A one-of-a-kind long-term data-collection effort, it is revealing how species and their interactions are responding to climate change, as well as shedding light on how species cope with habitat fragmentation.

In 1994, Hanski published the incidence-function model, which elegantly formulated the relationships between the area and isolation of a habitat patch and the likelihood of it being occupied by a species (I. Hanski *J. Anim. Ecol.* **63**, 151–162; 1994). This launched a new era of spatially explicit population models and was quickly adopted by ecologists.

Those of us who worked with Hanski will remember his sharp intellect and ceaseless enthusiasm for understanding nature. He was quick to adapt new methods and techniques. The Metapopulation Research Centre in Helsinki that Hanski established in 2000 consists of ecologists, evolutionary biologists, mathematicians, bioinformaticians and molecular biologists. Hanski also led Finland's first sequencing of an animal or plant genome: that of his Glanville butterfly, published in 2014 (V. Ahola *et al. Nature Commun.* **5**, 4737; 2014).

In 2003, nearly 25 years after completing his thesis, Hanski returned to his beloved dung beetles. He launched a project in Madagascar to study the evolutionary biology of the island's diverse endemic species of dung beetle and how these ecologically crucial communities respond to habitat loss. He led a series of excursions to Madagascar to work with local students, his team from Finland and his family. These trips became legendary, both for their scientific value and for the camaraderie he fostered.

Among numerous honours, Hanski was awarded ecology's top gong, the Crafoord Prize in Biosciences, in 2011. And despite his hectic schedule, he always prioritized public engagement. In Finland, Hanski was known for his views on conservation, in particular the protection of old-growth forests. Hanski was also a powerful advocate for basic research, criticizing science policy that demanded immediate economic benefit and arguing that such short-sighted aims threaten the fundamental process by which knowledge is generated.

Training young scientists was a top priority. Always keen to discuss ideas and offer feedback on manuscripts, Hanski became a co-author only on papers on which he felt he had made a significant intellectual contribution. This policy promoted the independence of early-career scientists working in the Metapopulation Research Centre. He loved to debate, and challenged everyone — regardless of their career stage — to discuss topics ranging from science to society.

Ilkka lived in Helsinki with his wife Eva Furman, a professor of environmental policy, and three children. He was dedicated to his family, a quality that resulted in a family-friendly working environment in the centre. After being diagnosed with cancer in 2014, Ilkka, with typical determination, completed projects closest to his heart — notably a book, *Messages from Islands: A Global Biodiversity Tour*, to be published in December.

Ilkka had so much more to give and he touched so many in the ecology community. His death leaves a gap that won't be filled. ■

**Anna-Liisa Laine** is professor of plant ecology at the University of Helsinki, Finland. Ilkka Hanski supervised her PhD between 2001 and 2005, and she joined the Metapopulation Research Centre as a group leader in 2010.  
e-mail: [anna-liisa.laine@helsinki.fi](mailto:anna-liisa.laine@helsinki.fi)

MARKUS MARCETIC



## HOST–MICROBE INTERACTION

# Rules of the game for microbiota

Are the dynamics of our microbial communities unique to us or does everyone's microbiota follow the same rules? The emerging insights into this question could be of relevance to health and disease. [SEE LETTER P.259](#)

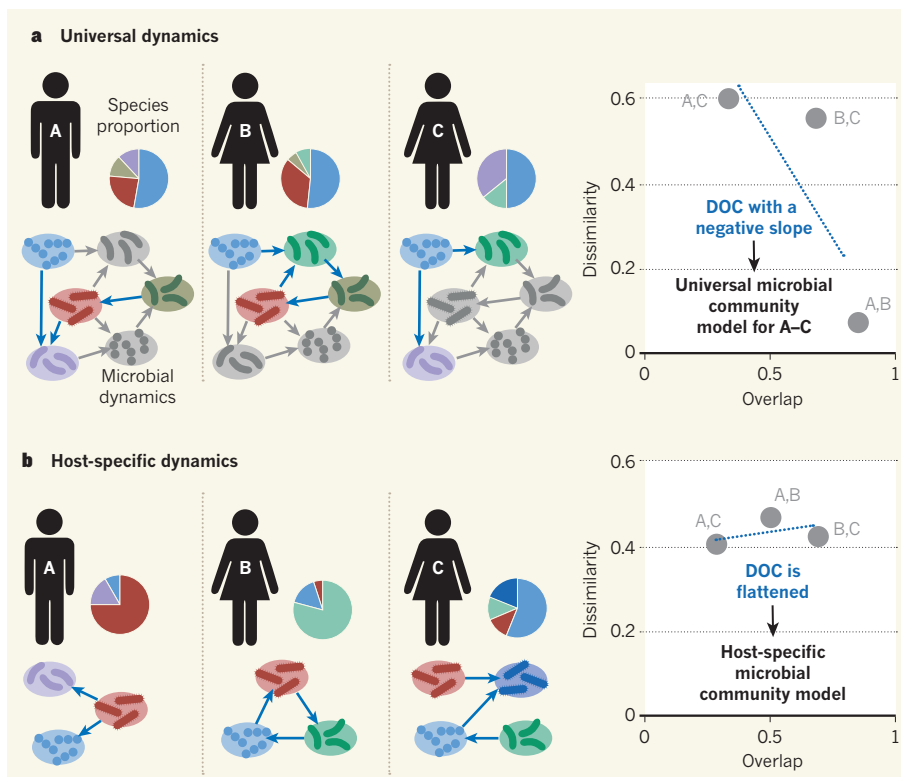
KAROLINE FAUST & JEROEN RAES

The composition of a body part's microbial community can differ substantially from one person to the next<sup>1–5</sup>. This is due to both host pressures and the dynamic behaviour of the microbes themselves. Understanding whether these interactions are consistent across hosts or whether each individual's microbiota follows its own rules has big implications. If the dynamics of an organ's microbial community are universal, we can use them to predict effective interventions for modulating the microbiota. If, however, microbial dynamics are host-specific, interventions must be designed separately for each person. Bashan *et al.*<sup>6</sup> address this issue using a new approach and report their intriguing observations on page 259.

To find out whether community dynamics are universal, ideally we should study long and densely sampled time series from many individuals with different traits and backgrounds. Models of microbial communities should then be fitted to the varying proportions of microbial species, which may become challenging when going beyond the most dominant groups of species. Such large temporal data sets are currently gravely lacking.

Bashan and colleagues devised an indirect method to address the question of universality. They measured two independent aspects of community similarity: overlap, which compares species assemblies by quantifying the proportion of shared species; and dissimilarity, which assesses the difference in abundance profiles of the shared species between individuals. The dissimilarity is then plotted against the overlap for all sample pairs to create a dissimilarity–overlap curve (DOC). If microbiota dynamics are truly universal (host-independent), then having the same species present should lead to the same relative proportion of those species, because they would dynamically influence each other in the same way. Consequently, a larger proportion of shared species should increase the community similarity and result in the tell-tale negative slope of the DOC (Fig. 1).

The authors tested their method by simulating microbial communities computationally using what is known as the generalized Lotka–Volterra model<sup>7</sup>, to generate communities with the same and with different



**Figure 1 | Learning from similarities and differences in our microbes.** To test whether microbial communities within a specific body part have the same underlying dynamics across individuals, Bashan *et al.*<sup>6</sup> used a method known as the dissimilarity–overlap curve (DOC). **a**, If microbial community dynamics are universal between individuals (A–C), the presence of the same species (species represented by coloured nodes; grey nodes represent absent species) should also lead to similar species proportions and a negative DOC slope. Consequently, a single model can be used to predict microbiota behaviour. **b**, If the community dynamics are host-specific, the presence of the same species does not lead to similar proportions and the DOC is flat. This necessitates the development of personalized models.

dynamics as positive and negative controls. In addition, they showed that randomizing data by shuffling microbial species across samples also removes the negative slope. These simulations confirm that the DOC detects universal dynamics and flattens in the absence of such dynamics. The curve even identifies strongly interacting species.

Most notably, the team detected negative slopes for the oral and gut communities in several human–microbiome data sets, including those of the Human Microbiome Project<sup>3</sup> and two human–gut time series<sup>8,9</sup>. However, the skin microbiota displayed weakly negative or flat DOCs in some cases, suggesting that

the microbial dynamics in the skin are host-specific at certain sites. Another interesting finding was that the DOC for the gut microbiota of people recurrently infected with the bacterial pathogen *Clostridium difficile*<sup>10</sup> is flat, but gains a negative slope after faecal transplantation from people who have not been infected.

If the assumptions hold, the consistent negative slopes observed for the healthy cohorts and for people treated after infection with *C. difficile* point to universal gut microbial dynamics. This is good news for all modelling efforts aiming to predict the behaviour of the gut microbiota during interventions or in disease. It means that when parameters such as growth rates and

interactions are determined for the gut microbial community of one healthy human, they are also valid for those of other individuals. Thus, the knowledge of such parameters can be combined across different studies and could, in the long term, allow a detailed, common microbial community model to be developed.

The DOC method has all the hallmarks of a powerful analytical tool. It is easy to implement, addresses a crucial question and may inspire applications beyond its intended use.

But, like all analyses, it makes a couple of assumptions — that the microbiota are in a steady state, and that having the same steady state implies that microbiota are governed by the same dynamics. The second assumption is the more risky: microbiota may end up in similar steady states not because of their intrinsic dynamics, but because of a strong environmental pressure that selects for a particular set of species. The authors rule out obvious host parameters such as diet, weight, age, race and transit time through the gut (measured by stool consistency) that may shape gut microbial communities. However, they do not account for all factors that may conceivably influence the gut microbiota<sup>2</sup>, and so cannot provide an entirely conclusive answer regarding the universality of the gut's microbial community dynamics.

The value of this work lies primarily in the importance of the question asked, the originality of the approach and the fact that it could spur a whole range of microbiome research. We expect it to spark fruitful discussions and lead to fresh ideas for analyses and experiments. For instance, it might be plausible to set up an artificial community under controlled conditions within a chemostat and then develop and define a model that describes its dynamics reasonably well. One could then compare the steady states reached by different subsets of the community to directly test the second assumption. If universal dynamics are confirmed, modelling efforts have a better chance of leading to more-effective clinical interventions. Bashan and colleagues' paper gives a glimpse of the deeper insights to be gained once we overcome the hurdles of controlled, high-throughput microbial community cultivation and manipulation. ■

**Karoline Faust and Jeroen Raes** are in the Department of Microbiology and Immunology, KU Leuven–University of Leuven, and in the VIB, Center for the Biology of Disease, Leuven 3000, Belgium. **K.F.** is also in the Microbiology Unit, Faculty of Sciences and Bioengineering Sciences, Vrije Universiteit Brussel.  
e-mail: jeroen.raes@kuleuven.be

6. Bashan, A. *et al. Nature* **534**, 259–262 (2016).
7. Stein, R. R. *et al. PLoS Comput. Biol.* **9**, e1003388 (2013).
8. Caporaso, J. G. *et al. Genome Biol.* **12**, R50 (2011).

9. David, L. A. *et al. Genome Biol.* **15**, R89 (2014).
10. Youngster, I. *et al. Clin. Infect. Dis.* **58**, 1515–1522 (2014).

## GEOCHEMISTRY

# Hydrogen and oxygen in the deep Earth

**The finding that an unusual iron oxide forms at extremely high pressures suggests that hydrogen and oxygen — two elements that strongly influence Earth's evolution — are generated in the mantle. SEE LETTER P.241**

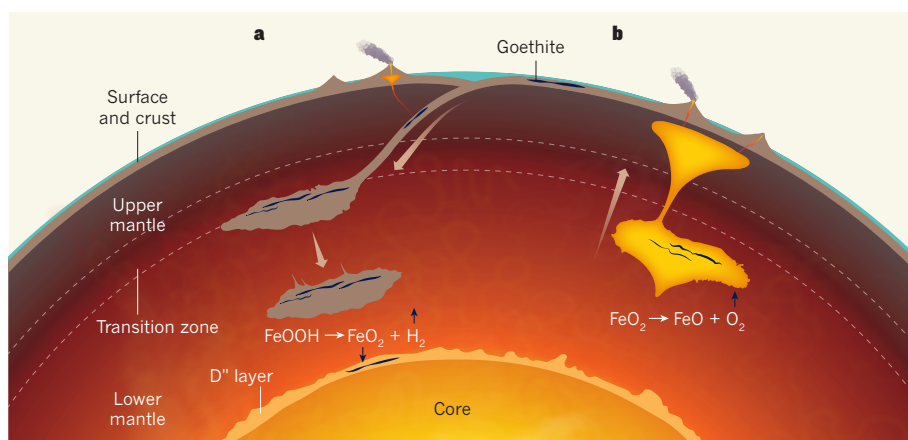
TAKEHIKO YAGI

Hydrogen greatly affects the properties of many materials. It is thought that most of the hydrogen in modern Earth is in water molecules, many of which are found in water-bearing minerals. It is therefore crucial to understand the stability and circulation of such hydrous minerals in Earth's interior, and this need has led to numerous studies of hydrous minerals under high-pressure and high-temperature conditions. In this issue, Hu *et al.*<sup>1</sup> (page 241) cast fresh light on the hydrogen-circulation issue. They report that an oxygen-rich iron oxide,  $\text{FeO}_2$ , is stabilized at pressures greater than about 76 gigapascals, and that this material might enable previously unknown hydrogen and oxygen cycles to occur in Earth's mantle.

Earth's core is mainly made of metallic iron, whereas the major minerals in the upper

mantle contain mostly ferrous iron ( $\text{Fe}^{2+}$ ). The most abundant form of iron on Earth's surface is haematite ( $\text{Fe}_2\text{O}_3$ ), which contains ferric iron ( $\text{Fe}^{3+}$ ) and is the main constituent of iron ore. Most of this ferric iron is thought to have formed by the oxidation of ferrous or metallic iron by the modern, oxygen-rich atmosphere.

On the basis of the distribution of ferric, ferrous and metallic iron from the surface to the core, it is thought that Earth's redox state becomes increasingly reducing with depth, so that the amount of ferric iron in the lower mantle would be limited. High-pressure laboratory experiments<sup>2,3</sup> revealed that, when olivine,  $(\text{Mg}, \text{Fe}^{2+})_2\text{SiO}_4$  (the most abundant mineral in the upper mantle) is subjected to conditions corresponding to those of the lower mantle, it changes into a mixture of two other minerals, bridgmanite,  $(\text{Mg}, \text{Fe}^{2+})\text{SiO}_3$ , and ferropericlase,  $(\text{Mg}, \text{Fe}^{2+})\text{O}$ . However, aluminium ions are also found in the mantle. When these are added,



**Figure 1 | Proposed source of hydrogen and oxygen in the lower mantle.** **a**, Descending slabs of Earth's crust can be carried to the transition zone between the upper and lower mantle, where they are heated until dense minerals form. The dense material then sinks to the bottom of the lower mantle. Hu *et al.*<sup>1</sup> suggest that when the mineral goethite ( $\text{FeOOH}$ , commonly formed by the reaction of the mineral haematite and water on Earth's surface) is carried to the mantle by a slab, an oxygen-rich iron oxide ( $\text{FeO}_2$ ) and hydrogen would form at depths greater than 1,800 kilometres. The dense  $\text{FeO}_2$  would sink to the bottom of the lower mantle, and might help to explain the structural complexity of the  $\text{D}''$  layer, which lies close to the core–mantle boundary. The highly mobile hydrogen would spread upwards. **b**, If the  $\text{FeO}_2$ -containing material is lifted by motion in the lower mantle, it will break down and release oxygen at depths of less than 1,500 km. (Adapted from a graphic by Jun Tsuchiya.)

1. Arumugam, M. *et al. Nature* **473**, 174–180 (2011).
2. Falony, G. *et al. Science* **352**, 560–564 (2016).
3. The Human Microbiome Project Consortium. *Nature* **486**, 207–214 (2012).
4. Zhernakova, A. *et al. Science* **352**, 565–569 (2016).
5. Costello, E. K. *et al. Science* **326**, 1694–1697 (2009).



bridgmanite containing a large amount of ferric iron forms, together with ferropericlase and some metallic iron<sup>4</sup>. More than 60% of the total iron in bridgmanite can be ferric iron.

Hu *et al.* now add to this picture by investigating what happens when haematite is compressed in oxygen and heated to generate the pressure and temperature conditions that correspond to the deep lower mantle (78 GPa and 1,800 kelvin). The authors used X-ray diffraction to study the sample, but the diffraction patterns obtained were quite 'spotty' and the sample existed neither as a powder nor as a single crystal. In such cases, the crystal structures of materials cannot be determined in detail. The researchers therefore used a method called multigrain crystallography<sup>5</sup> to analyse the spotty patterns, and concluded that the sample is an aggregate of at least 33 single crystals in which the haematite has changed into FeO<sub>2</sub>, an iron oxide that has the same atomic structure as pyrite (FeS<sub>2</sub>).

This finding might suggest that Fe<sup>4+</sup> — normally a metastable form of iron — had formed under the extreme experimental conditions, and that its charge is balanced by two O<sup>2-</sup> ions. However, Hu *et al.* found that the oxygen–oxygen (O–O) bond in FeO<sub>2</sub> is only 1.937 ångströms in length; by comparison, the ionic radius of O<sup>2-</sup> is 1.40 Å (ref. 6), corresponding to an O–O bond length of 2.8 Å or greater. The observed bond length does, however, agree with the typical O–O bond length for a peroxide ion (O<sub>2</sub><sup>2-</sup>). If the sample contains peroxide ions, then the iron must be ferrous, to balance the charge of those ions; in other words, the iron has been reduced from Fe<sup>3+</sup> in haematite to Fe<sup>2+</sup>. Such a reaction is possible only at very high pressures, because the FeO<sub>2</sub> has a smaller volume than a mixture of haematite and oxygen, and the smaller volume becomes energetically favourable under extreme high pressures.

Hu and colleagues went on to show that the mineral goethite, FeOOH, also forms FeO<sub>2</sub> at 2,050 K and a pressure of 92 GPa by releasing hydrogen. Goethite commonly forms from the reaction of haematite and water at Earth's surface. The authors further demonstrated that the FeO<sub>2</sub> formed in this way becomes unstable, and probably breaks down into ferrous oxide (FeO) and oxygen when the pressure is reduced.

These findings present new possibilities for how hydrogen and oxygen form and circulate inside Earth. When goethite (or a mixture of haematite and water) is carried deep into the lower mantle by subduction processes, then hydrogen and FeO<sub>2</sub> are formed (Fig. 1). Hydrogen is extremely mobile and will spread upwards, eventually escaping into space, whereas heavy FeO<sub>2</sub> will sink to the bottom of the lower mantle. But if the FeO<sub>2</sub> is lifted to the upper part of the lower mantle by, for example, an upwelling plume of hot rock, it will become unstable and release oxygen gas on the way. This means that large amounts of hydrogen

and oxygen might occasionally be produced in the lower mantle.

No such possibility has previously been considered. As the authors claim, this process could have acted as an additional or alternative oxygen source for the Great Oxidation Events — the periods in Earth's history when the atmosphere became oxygenated. Until now, it was thought that the oxygen was supplied by biological activity alone.

If hydrogen is released by the subduction of goethite, how will it behave at great depths? Unfortunately, hydrogen is invisible to X-rays and to electron microscopy, which makes its behaviour difficult to study at the atomic scale. Neutron diffraction is a powerful probe for directly observing hydrogen, and a facility<sup>7,8</sup> that enables this technique to be used at high pressures and temperatures has successfully tracked the movement of hydrogen within materials<sup>9</sup>. Such techniques could be used to study the behaviour of hydrogen in Earth's interior.

And what is the fate of FeO<sub>2</sub> when it sinks to the bottom of the lower mantle? Just like the finding<sup>10</sup> that bridgmanite adopts an

unexpected dense phase at pressures greater than 120 GPa, Hu and colleagues' work suggests explanations for the structural complexity of the region called the D'' layer near the core–mantle boundary. Further studies are required to address this issue, and to work out how hydrogen and oxygen circulate in the deep Earth. ■

**Takehiko Yagi** is in the Geochemical Research Centre, University of Tokyo, Tokyo 113-0033, Japan.

e-mail: yagi@eqchem.s.u-tokyo.ac.jp

1. Hu, Q. *et al.* *Nature* **534**, 241–244 (2016).
2. Liu, L.-G. *Phys. Earth Planet. Inter.* **11**, 289–298 (1976).
3. Ito, E. & Takahashi, E. *J. Geophys. Res. Solid Earth* **94**, 10637–10646 (1989).
4. Frost, D. J. *et al.* *Nature* **428**, 409–412 (2004).
5. Sørensen, H. O. *et al.* *Z. Kristallogr.* **227**, 63–78 (2012).
6. Shannon R. D. & Prewitt C. T. *Acta Crystallogr. B* **25**, 925–946 (1969).
7. Hattori H. *et al.* *Nucl. Instrum. Methods Phys. Res. A* **780**, 55–67 (2015).
8. Sano-Furukawa A. *et al.* *Rev. Sci. Instrum.* **85**, 113905 (2014).
9. Machida, A. *et al.* *Nature Commun.* **5**, 5063 (2014).
10. Murakami, M., Hirose, K., Kawamura, K., Sata, N. & Ohishi, Y. *Science* **304**, 855–858 (2004).

#### EVOLUTIONARY BIOLOGY

## To mimicry and back again

**Deadly coral snakes warn predators through striking red–black banding. New data confirm that many harmless snakes have evolved to resemble coral snakes, and suggest that the evolution of this Batesian mimicry is not always a one–way street.**

DAVID W. PFENNIG

Many species that are dangerous or unpleasant to eat have evolved conspicuous signals that warn predators to avoid them. Not surprisingly, many other species that are edible to predators, from birds and butterflies to salamanders and sea slugs, have evolved to resemble these inedible species<sup>1</sup>. By doing so, the 'mimics' receive protection from predation, just like their 'models'. This phenomenon is known as Batesian mimicry after the explorer and naturalist Henry Walter Bates, who first described it<sup>2</sup>. Batesian mimicry has long fascinated evolutionary biologists, and it is widely used to illustrate the power of natural selection to produce remarkable adaptation<sup>1</sup>. Yet we still do not know how common Batesian mimicry is, what its role is in evolutionary diversification, nor whether it can be reversed. Writing in *Nature Communications*, Davis Rabosky *et al.*<sup>3</sup> present findings on mimicry of coral snakes that go a long way towards answering these questions.

In 1867, the naturalist Alfred Russel Wallace

suggested that the striking resemblance between deadly coral snakes and numerous harmless species of red–black-banded (RBB) snakes reflected Batesian mimicry<sup>4</sup> (Fig. 1). However, whether coral-snake mimicry actually occurs has been questioned ever since, primarily because of the (presumed) non-concordance in the geographical distributions and abundances of coral snakes and their mimics. Several studies have attempted to address this issue<sup>5</sup>; most notably, it has been demonstrated that predators avoid artificial snakes that have RBB patterns<sup>6</sup> but only in geographical regions where coral snakes occur<sup>7</sup>, exactly as predicted by the mimicry hypothesis.

Davis Rabosky and colleagues focus on this system, but present a more comprehensive study than these earlier investigations. By integrating colour-pattern, distribution and phylogenetic data across all 'New World' species of snake, they show that evolutionary shifts to RBB patterns in coral snakes and numerous non-venomous species are highly correlated in space and time. Indeed, they find that every origin of the RBB pattern in non-venomous

## EVOLUTIONARY BIOLOGY

# To mimicry and back again

**Deadly coral snakes warn predators through striking red–black banding. New data confirm that many harmless snakes have evolved to resemble coral snakes, and suggest that the evolution of this Batesian mimicry is not always a one-way street.**

DAVID W. PFENNIG

Many species that are dangerous or unpleasant to eat have evolved conspicuous signals that warn predators to avoid them. Not surprisingly, many other species that are edible to predators, from birds and butterflies to salamanders and sea slugs, have evolved to resemble these inedible species<sup>1</sup>. By doing so, the ‘mimics’ receive protection from predation, just like their ‘models’. This phenomenon is known as Batesian mimicry after the explorer and naturalist Henry Walter Bates, who first described it<sup>2</sup>. Batesian mimicry has long fascinated evolutionary biologists, and it is widely used to illustrate the power of natural selection to produce remarkable adaptation<sup>1</sup>. Yet we still do not know how common Batesian mimicry is, what its role is in evolutionary diversification, nor whether it can be reversed. Writing in *Nature Communications*, Davis Rabosky *et al.*<sup>3</sup> present findings on mimicry of coral snakes that go a long way towards answering these questions.

In 1867, the naturalist Alfred Russel Wallace

suggested that the striking resemblance between deadly coral snakes and numerous harmless species of red–black-banded (RBB) snakes reflected Batesian mimicry<sup>4</sup> (Fig. 1). However, whether coral-snake mimicry actually occurs has been questioned ever since, primarily because of the (presumed) non-concordance in the geographical distributions and abundances of coral snakes and their mimics. Several studies have attempted to address this issue<sup>5</sup>; most notably, it has been demonstrated that predators avoid artificial snakes that have RBB patterns<sup>6</sup> but only in geographical regions where coral snakes occur<sup>7</sup>, exactly as predicted by the mimicry hypothesis.

Davis Rabosky and colleagues focus on this system, but present a more comprehensive study than these earlier investigations. By integrating colour-pattern, distribution and phylogenetic data across all ‘New World’ species of snake, they show that evolutionary shifts to RBB patterns in coral snakes and numerous non-venomous species are highly correlated in space and time. Indeed, they find that every origin of the RBB pattern in non-venomous

snakes occurred only after that particular lineage and coral snakes were present together in the New World. Thus, in every case, the warning signal arose first in the model, then in the mimic, which is a key prediction of Batesian-mimicry theory. These data should therefore lay to rest any doubts about whether coral-snake mimicry does occur.

The authors’ work also shows that coral-snake diversity strongly predicts (and substantially increases) the number of mimic species in a given geographical area. Indeed, their data suggest that the ‘mimicry excess’ problem is even greater than has been historically assumed, with up to six times more mimetic than model species present in a given locality, and many more than would be expected if RBB snakes were distributed randomly across the New World. These data are at odds with the long-standing theoretical expectation that Batesian mimics should be rarer than their toxic models. However, this expectation might not apply with a highly toxic model, such as the coral snake. When the model is highly toxic, the fitness costs of mistakenly attacking it would probably be so severe that predators would be under strong selection to avoid such a model (and any lookalikes), even if the model is rare<sup>8</sup>.

Another advance that stems from this work is the authors’ proposal that mimicry might not represent an evolutionary end point. In particular, their data suggest that not only have evolutionary transitions between cryptic (non-mimetic) patterns and RBB (mimetic) patterns occurred frequently in non-venomous snakes, but so also have transitions between mimetic patterns and cryptic patterns. Most of these losses of mimicry occurred in the tropics, where coral snakes are continuously distributed,

DONALD B. SHEPPARD



**Figure 1 | Protective imitation.** Many species of harmless snake, such as the false coral snake, *Erythrolamprus aesculapii* (left), have evolved the red–black-banded (RBB) colour pattern of highly venomous coral snakes, such as the Brazilian coral snake, *Micrurus brasiliensis* (right). Davis Rabosky *et al.*<sup>3</sup> show that this RBB pattern has evolved in multiple lineages of non-venomous snakes, but only after each lineage and coral snakes were present together in the New World, supporting the long-standing Batesian-mimicry hypothesis<sup>2</sup>.



suggesting that these losses occurred even among species that live in the same area as coral snakes. This is an intriguing conclusion. Generally, mimicry has been viewed as a one-way street; it is not clear why a species should ever lose mimicry once it has evolved it, particularly if their model is still around.

This suggestion will no doubt motivate further studies to determine how transitions between mimicry and cryptic patterning occur. Evolutionary biologists have long debated whether Batesian mimicry could evolve through a gradual process of incremental evolution<sup>9</sup>, and many of these arguments should apply equally to its loss. In particular, it is unclear how a population could transition from an ancestral cryptic colour pattern

to a derived mimetic one (or vice versa) if the population must pass through a phase in which it expresses a colour pattern that is intermediate between these two extremes. Such an intermediate colour pattern would be expected to be disfavoured, because it should fail to receive the fitness benefits of either cryptic patterning or mimicry.

Batesian mimicry has been called<sup>9</sup> “the greatest post-Darwinian application of Natural Selection”. Davis Rabosky and colleagues’ study has settled some questions regarding the specific example of coral-snake mimicry, and it opens the door to answering several others. ■

**David W. Pfennig** is in the Department of Biology, University of North Carolina, Chapel

Hill, North Carolina 27599, USA.

e-mail: [dpfennig@unc.edu](mailto:dpfennig@unc.edu)

1. Stevens, M. *Cheats and Deceits: How Animals and Plants Exploit and Mislead* (Oxford Univ. Press, 2016).
2. Bates, H. W. *Trans. Linn. Soc. Lond.* **23**, 495–566 (1862).
3. Davis Rabosky, A. R. *et al. Nature Commun.* **7**, 11484 (2016).
4. Wallace, A. R. *Westminster Foreign Q. Rev.* **32** (1), 1–43 (1867).
5. Greene, H. W. & McDiarmid, R. W. *Science* **213**, 1207–1212 (1981).
6. Brodie, E. D. III & Janzen, F. J. *Funct. Ecol.* **9**, 186–190 (1995).
7. Pfennig, D. W., Harcombe, W. R. & Pfennig, K. S. *Nature* **410**, 323 (2001).
8. Lindström, L., Alatalo, R. V. & Mappes, J. *Proc. R. Soc. B* **264**, 149–153 (1997).
9. Fisher, R. A. *The Genetical Theory of Natural Selection* (Clarendon, 1930).



**Figure 1 | Protective imitation.** Many species of harmless snake, such as the false coral snake, *Erythrolamprus aesculapii* (left), have evolved the red-black-banded (RBB) colour pattern of highly venomous coral snakes, such as the Brazilian coral snake, *Micrurus brasiliensis* (right). Davis Rabosky *et al.*<sup>3</sup> show that this RBB pattern has evolved in multiple lineages of non-venomous snakes, but only after each lineage and coral snakes were present together in the New World, supporting the long-standing Batesian-mimicry hypothesis<sup>2</sup>.

snakes occurred only after that particular lineage and coral snakes were present together in the New World. Thus, in every case, the warning signal arose first in the model, then in the mimic, which is a key prediction of Batesian-mimicry theory. These data should therefore lay to rest any doubts about whether coral-snake mimicry does occur.

The authors' work also shows that coral-snake diversity strongly predicts (and substantially increases) the number of mimic species in a given geographical area. Indeed, their data suggest that the 'mimicry excess' problem is even greater than has been historically assumed, with up to six times more mimetic than model species present in a given locality, and many more than would be expected if RBB snakes were distributed randomly across the New World. These data are at odds with the long-standing theoretical expectation that Batesian mimics should be rarer than their toxic models. However, this expectation might not apply with a highly toxic model, such as the coral snake. When the model is highly toxic, the fitness costs of mistakenly attacking it would probably be so severe that predators would be under strong selection to avoid such a model (and any lookalikes), even if the model is rare<sup>8</sup>.

Another advance that stems from this work is the authors' proposal that mimicry might not represent an evolutionary end point. In particular, their data suggest that not only have evolutionary transitions between cryptic (non-mimetic) patterns and RBB (mimetic) patterns occurred frequently in non-venomous snakes, but so also have transitions between mimetic patterns and cryptic patterns. Most of these losses of mimicry occurred in the tropics, where coral snakes are continuously distributed, suggesting that these losses occurred even among species that live in the same area as coral snakes. This is an intriguing conclusion. Generally, mimicry has been viewed as a one-way street; it is not clear why a species should ever lose mimicry once it has evolved it, particularly if their model is still around.

This suggestion will no doubt motivate further studies to determine how transitions between mimicry and cryptic patterning occur. Evolutionary biologists have long debated whether Batesian mimicry could evolve through a gradual process of incremental evolution<sup>9</sup>, and many of these arguments should apply equally to its loss. In particular, it is unclear how a population could transition from an ancestral cryptic colour pattern to a derived mimetic one (or vice versa) if the population must pass through a phase in which it expresses a colour pattern that is intermediate between these two extremes. Such an intermediate colour pattern would be expected to be disfavoured, because it should fail to receive the fitness benefits of either cryptic patterning or mimicry.

Batesian mimicry has been called<sup>9</sup> "the greatest post-Darwinian application of Natural Selection". Davis Rabosky and colleagues' study has settled some questions regarding the

specific example of coral-snake mimicry, and it opens the door to answering several others. ■

David W. Pfennig is in the Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599, USA.

e-mail: dpfennig@unc.edu

1. Stevens, M. *Cheats and Deceits: How Animals and Plants Exploit and Mislead* (Oxford Univ. Press, 2016).
2. Bates, H. W. *Trans. Linn. Soc. Lond.* **23**, 495–566 (1862).
3. Davis Rabosky, A. R. *et al. Nature Commun.* **7**, 11484 (2016).
4. Wallace, A. R. *Westminster Foreign Q. Rev.* **32** (1), 1–43 (1867).
5. Greene, H. W. & McDiarmid, R. W. *Science* **213**, 1207–1212 (1981).
6. Brodie, E. D. III & Janzen, F. J. *Funct. Ecol.* **9**, 186–190 (1995).
7. Pfennig, D. W., Harcombe, W. R. & Pfennig, K. S. *Nature* **410**, 323 (2001).
8. Lindström, L., Alatalo, R. V. & Mappes, J. *Proc. R. Soc. B* **264**, 149–153 (1997).
9. Fisher, R. A. *The Genetical Theory of Natural Selection* (Clarendon, 1930).

This article was published online on 25 May 2016.

#### PHYSIOLOGY

## Microbial signals to the brain control weight

**The bacteria that inhabit the rodent gut promote insulin secretion and food intake by activating the parasympathetic nervous system — a hitherto unknown mode of action for this multifaceted microbiota. [SEE ARTICLE P.213](#)**

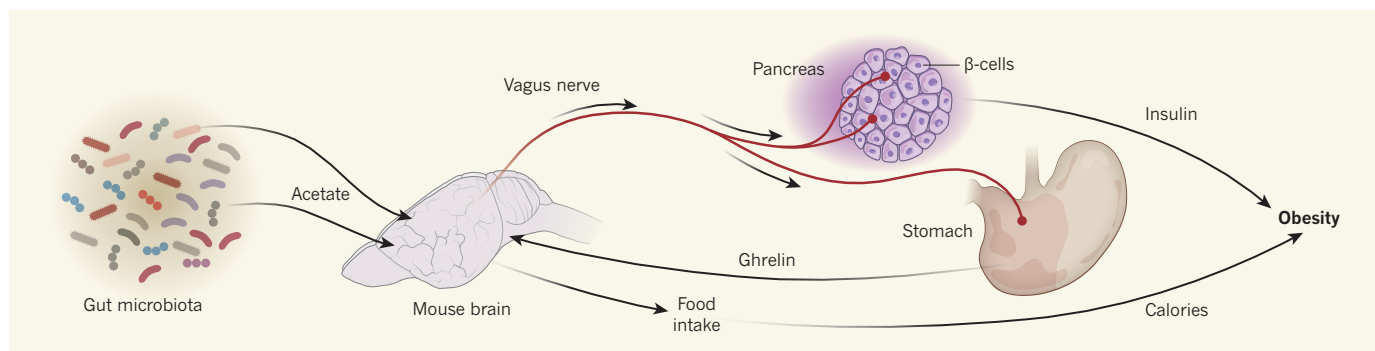
MIRKO TRAJKOVSKI & CLAES B. WOLLHEIM

**W**e live in symbiosis with trillions of bacteria that populate our intestines, known collectively as the gut microbiota. These microbes influence many physiological processes in our bodies, from gut and immune maintenance to neurological regulation<sup>1</sup>. On page 213 of this issue, Perry

*et al.*<sup>2</sup> highlight a previously unknown role for the gut microbiota in stimulating insulin secretion by signalling to the brain. Moreover, the authors report that these microbes influence appetite, providing a hint as to how the microbiota might provoke obesity.

Mammals have evolved several responses to energy scarcity. As a result of these adaptations, obesity can arise in conditions of





**Figure 1 | A mechanism for microbiota-mediated weight gain.** Perry *et al.*<sup>2</sup> report that, in rodents, production of acetate molecules from dietary nutrients by the bacteria that colonize the gut (the microbiota) increases the brain's stimulation of the parasympathetic nervous system, which includes the vagus nerve. Signals from the vagus nerve trigger

secretion of the 'hunger hormone' ghrelin from the stomach, leading to increased food intake. The vagus nerve also potentiates glucose-stimulated insulin secretion from  $\beta$ -cells in the pancreas, promoting calorie storage and fat gain. In this way, the gut microbiota influences obesity.

constant food abundance. This response is mediated by the hormone insulin, which is secreted from pancreatic  $\beta$ -cells in response to increased blood-glucose levels. Insulin tightly controls energy balance by enhancing cellular lipid synthesis and glucose uptake, causing calorie storage.

Investigating the effects of a high-fat diet, Perry and colleagues found that production and turnover of the short-chain fatty acid (SCFA) acetate was markedly increased in rats on a high-fat diet compared with animals fed a normal diet. Moreover, infusing the stomachs of rats on a normal diet with acetate for ten days increased glucose-stimulated insulin secretion (GSIS).

Although glucose is the main stimulus for insulin secretion, the process is also under the control of the parasympathetic nervous system<sup>3</sup> — the part of the central nervous system that stimulates 'rest-and-digest' and 'feed-and-breed' processes. Parasympathetic activity is largely mediated by the vagus nerve, which sends motor inputs to many organs and is responsible for slowing heart rate, and for regulating gastrointestinal movement and the digestion of food, in addition to enhancing insulin secretion<sup>4</sup>. Perry *et al.* demonstrated that the ability of acetate infusion to increase GSIS could be blocked by administering the parasympathetic blocker molecules atropine or methylatropine, or by surgically severing one or more of the branches of the vagus nerve that connects to the gut. These results indicate that an acetate-induced increase in GSIS is controlled by the parasympathetic nervous system.

Further supporting the role of the parasympathetic nervous system in acetate-mediated GSIS, the authors demonstrated that acetate could not stimulate insulin secretion from isolated  $\beta$ -cell-containing pancreatic islets *in vitro*. This is consistent with some, but not all, previous investigations into a direct effect of acetate on  $\beta$ -cells (for a review, see ref. 5). Acetate administration into either the brain's ventricular system or a vertical column

of grey matter embedded in the brainstem — both of which feed into the parasympathetic nervous system — increased GSIS, again highlighting the central-nervous effects of acetate.

Next, Perry *et al.* investigated the effects of increased acetate turnover on appetite. A chronic increase in acetate turnover promoted a constant drive to eat, known as hyperphagia, probably mediated by the 'hunger hormone' ghrelin — levels of which were elevated in the hyperphagic rats compared with controls. The hyperphagic rats developed obesity, probably owing to a combination of increased secretion of ghrelin and insulin.

Because SCFAs are products of bacterial fermentation, Perry and co-workers investigated the role of the gut microbiota in acetate turnover. The gut microbiota co-develops with the host and modulates whole-body metabolism by affecting energy balance<sup>6–9</sup>. The authors transplanted faecal matter from donor rats on a normal or high-fat diet into recipients on the opposing diet, and found that the acetate-turnover rate, faecal acetate levels and GSIS levels from the donor group were transferred to the recipients, implying that it is changes in the microbiota that regulate these factors. Furthermore, conditions of microbiota depletion (seen in germ-free mice, which lack a microbiota, or in rats treated with antibiotics) completely suppressed acetate turnover and decreased ghrelin levels compared to control mice — changes that were associated with two- and fivefold lower skeletal-muscle fat content, respectively.

These data suggest a mechanistic link between the onset of obesity and the gut microbiota. The microbiota-mediated increase in acetate turnover that occurs during exposure to a high-calorie diet might mediate a feedback loop between the gut microbiota and parasympathetic nervous system, promoting hyperphagia owing to increased ghrelin secretion, and increased energy storage as fat owing to increased GSIS (Fig. 1). However, this mechanism does not explain the observation<sup>10</sup> that microbiota-depleted mice do not show suppressed food intake. It is also intriguing that

supplementation of the diets of rats with two other SCFAs, butyrate and propionate, improves host physiology and glucose metabolism, which in the case of propionate seems to be mediated by vagus-nerve stimulation by the peripheral nervous system<sup>11</sup>. This might indicate that the site of stimulation — central or peripheral — is relevant for SCFA-mediated effects in the parasympathetic nervous system, and points to the need for further exploration of the general role of SCFAs in regulating obesity.

For instance, follow-up work could address whether the effects in the brain are mediated by the SCFA receptor proteins FFA2 and FFA3, and clarify the controversy<sup>5</sup> regarding the direct effects of acetate on the  $\beta$ -cells. In addition, transplantation of the microbiota from rodents on a high-fat diet or from humans who are obese to germ-free rodents fed a normal diet could allow researchers to further test for a causal link between specific obesity-associated changes brought on by microbiotic acetate production and the development of metabolic syndrome (which involves obesity, insulin resistance, abnormal lipid levels in the blood and glucose intolerance). Analysis of how the genomes of the microbiota collectively change in rodents on a high-fat diet would allow researchers to identify acetate-producing microbes and to investigate their importance in the progression of diet-induced obesity.

Clinical trials<sup>3</sup> have shown that vagus-nerve blockade by electrodes can help to reduce body weight and improve blood-glucose control in people with obesity. Moreover, specific antimicrobials and phage therapies<sup>12</sup>, as well as faecal or bacterial transfers, have attracted renewed interest in the past few years as potential tools to treat antibiotic-resistant enteritis (inflammation of the intestine) and ulcerative colitis<sup>13</sup> (long-term inflammation of the colon and rectum). In the context of the increased global prevalence of obesity, Perry and colleagues' study might inform the development of such strategies for suppressing acetate or acetate-producing microbes as a means to treat obesity and diabetes. ■

**Mirko Trajkovski and Claes B. Wollheim**  
are in the Department of Cell Physiology and Metabolism, Centre Médical Universitaire, Faculty of Medicine, University of Geneva, 1211 Geneva, Switzerland.  
e-mails: mirko.trajkovski@unige.ch;  
claes.wollheim@unige.ch

1. Sommer, F. & Bäckhed, F. *Nature Rev. Microbiol.* **11**, 227–238 (2013).

2. Perry, R. J. *et al. Nature* **534**, 213–217 (2016).  
3. de Lartigue, G. J. *Physiol. (Lond.)* <http://dx.doi.org/10.1113/JP271538> (2016).  
4. Bereiter, D. A., Berthoud, H. R., Brunsmann, M. & Jeanrenaud, B. *Am. J. Physiol.* **241**, E22–E27 (1981).  
5. Priyadarshini, M., Wicksteed, B., Schiltz, G. E., Gilchrist, A. & Layden, B. T. *Trends Endocrinol. Metab.* <http://dx.doi.org/10.1016/j.tem.2016.03.011> (2016).  
6. Turnbaugh, P. J. *et al. Nature* **444**, 1027–1031 (2006).  
7. Koren, O. *et al. Cell* **150**, 470–480 (2012).

8. Chevalier, C. *et al. Cell* **163**, 1360–1374 (2015).  
9. Bäckhed, F. *et al. Proc. Natl Acad. Sci. USA* **101**, 15718–15723 (2004).  
10. Suárez-Zamorano, N. *et al. Nature Med.* **21**, 1497–1501 (2015).  
11. De Vadder, F. *et al. Cell* **156**, 84–96 (2014).  
12. Reyes, A., Semenov, N. P., Whiteson, K., Rohwer, F. & Gordon, J. I. *Nature Rev. Microbiol.* **10**, 607–617 (2012).  
13. Petrof, E. O. & Khoruts, A. *Gastroenterology* **146**, 1573–1582 (2014).

## CHEMISTRY

# No turning back for motorized molecules

**Two molecular motors have been developed that use chemical energy to drive rotational motion in a single direction. The findings bring the prospect of devices powered by such motors a tantalizing step closer. SEE LETTER P.235**

JONATHAN CLAYDEN

The conversion of chemical energy to mechanical motion drives movement in all living things, from bacteria to whales. An intricate array of molecular ratchets and motors allows cells to extract mechanical work from chemical reactions, for example to drive muscle contraction, or to twist the helical appendages that propel some bacteria. Two papers, one by Collins *et al.*<sup>1</sup> in *Nature Chemistry* and another by Wilson *et al.*<sup>2</sup> on page 235, report the design and construction of artificial molecular motors that achieve the same outcome using much simpler, purely synthetic structures. Both pieces of work show that chemical reagents can drive the unidirectional motion of one part of a molecule (the rotor) relative to another (the stator), and thus provide direct functional analogues of biological motors.

It is not easy to design a synthetic molecular motor<sup>3</sup>. As was pointed out nearly 20 years ago<sup>4,5</sup>, molecular motors are characterized by movement that must be more than random Brownian motion. Furthermore, angular momentum cannot be used to maintain a constant directionality on the molecular scale as it can in everyday electric motors. The thermodynamic landscape of a molecular system must be repeatedly altered to force concerted movement in a single direction, to prevent mere shuttling forwards and backwards between two states. The greatest successes in the field so far have used light energy to drive a molecular system away from equilibrium, followed by a directionally defined relaxation process; motors capable of megahertz rotational speeds have been designed and built using this approach<sup>6</sup>.

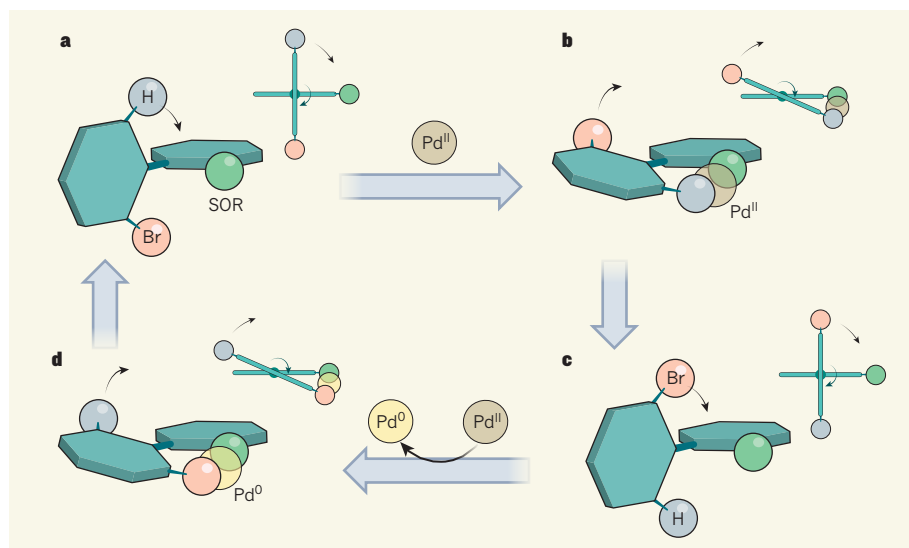
The two new motors both use chemical energy to drive rotation. Collins and colleagues' motor is remarkably simple in

conception. The rotor and stator are each a benzene ring, connected by a single bond that forms a rotatable axle. Systems of this sort can rotate freely about the axle, but rotation in the authors' motor is partly restricted by groups or atoms attached next to the bond that connects the two benzene rings.

Collins *et al.* added alternating sets of reagents to a solution containing their motor,

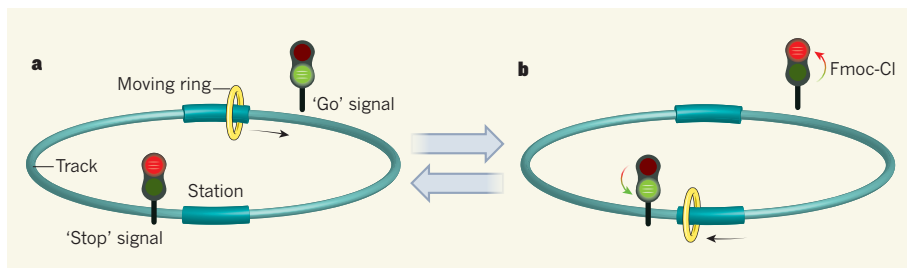
which allowed first one side of the rotor ring and then the other to slip past a sulfur-containing group (a sulfoxide; Fig. 1) bonded to the stator. The alternating reagents insert a palladium atom into a carbon–hydrogen (C–H) bond on one side of the rotor, and then into a carbon–bromine (C–Br) bond on the other. Palladium's affinity for the sulfur atom of the sulfoxide (SOR) group lets it form a bridge between the rings that lowers the energy barrier to rotation, allowing the rings to slip past one another. On its own, shuttling the metal between the C–H and C–Br bonds would simply cause random rotation clockwise or anticlockwise, but the chirality (handedness) of the sulfoxide group imparts directionality to the slippage mechanism, and so also to the rotation of the motor.

The alternating C–H and C–Br insertions needed to drive this process require palladium to be in the +2 and 0 oxidation states, respectively. This means that, in its current form, the motor cannot work autonomously, because different reaction conditions are needed to



**Figure 1 | A unidirectional molecular motor incorporating a rotating axle.** Collins *et al.*<sup>1</sup> report a system consisting of two benzene rings (green hexagons) connected by a single bond. One ring acts as a rotor, and has a hydrogen atom on one side and a bromine atom on the other. The other ring is a stator and has a sulfoxide group on one side and a fluorine atom on the other (fluorine atom not shown because it is not involved in the motor mechanism). The connecting bond acts as an axle. The rings are also viewed here from above, along the axis of the axle (top right in each panel). **a**, The system's rotation cycle begins with the rotor carrying the hydrogen atom to pass the sulfoxide. A palladium atom bridges the two rings. **c**, The rings then relax to the alternative perpendicular arrangement. **d**, Conversion of palladium(II) to palladium(0) allows the side of the rotor carrying the bromine atom to pass the sulfoxide group, and a palladium atom again bridges the rings. The cycle continues if reagents are added to toggle the palladium between the two oxidation states. Br, bromine; SOR, sulfoxide (where R is a benzene-ring-containing group); Pd, palladium.





**Figure 2 | A unidirectional molecular motor incorporating a ring travelling round a track.** **a**, Wilson *et al.*<sup>2</sup> report a system in which a small molecular ring is threaded onto a larger one (the track). The track has two 'stations' at which the ring can dock, and two molecular groups that act as signals and can be set to 'stop' or 'go'. When the ring is docked at a station, its proximity to the nearby signal forces that signal to stay in the 'go' position, allowing the ring to travel to the second station. **b**, When the ring docks at the second station, the first signal changes to 'stop', preventing reversal of the direction of travel, while the second signal switches to 'go', allowing the ring to carry on around the track to the first station. The signals are switched from 'go' to 'stop' using the reagent Fmoc-Cl.

shuttle the palladium between these states. However, metal redox processes can be driven electrochemically, raising the intriguing possibility that future versions of the motor could be electrically powered.

Wilson and colleagues' chemically powered motor overcomes the autonomy problem by using a different, more complex design than that of Collins and co-workers. In Wilson and colleagues' motor, a small ring is threaded onto a larger one (the track), and travels like a train around the track by constantly advancing in the same direction from one of two 'stations' to the other. Crucially, and in contrast to Collins' and co-workers' motor, only one set of reaction conditions is needed to drive the motor forward: the authors use a reactive 'fuel' known as Fmoc-Cl, which continuously breaks down to carbon dioxide and other by-products as the motor runs.

The small ring's progress is powered by a mechanism that channels random kinetic motion into movement in a single direction (Fig. 2). Immediately after each station, an unstable carbonate group (which becomes attached to the track by reaction with Fmoc-Cl) provides a 'stop' signal. If the carbonate group is removed, the signal switches to 'go'. The authors designed the chemistry of the system such that the signal switches from 'stop' to 'go' at a more or less constant rate at both stations, but changes from 'go' to 'stop' more rapidly after the small ring has passed through to the other station. The stop signal therefore tends to follow the train around the track, ensuring that forward motion is always faster than reverse.

The choice of Fmoc-Cl as the fuel is ingenious, because the chemical mechanisms that involve Fmoc-Cl in switching from 'stop' to 'go' and vice versa are different, which means that the rates of the switching steps can be independently controlled. The energy that powers the constant forward movement of the train is provided by the consumption of Fmoc-Cl, so the train keeps moving until all of the Fmoc-Cl has been consumed.

Wilson and colleagues' work constitutes

an important step in the construction of a chemically propelled, autonomous molecular device, but there is still a long way to go. The small ring typically takes 12 hours to travel around the track, and the Fmoc-Cl fuel is used rather inefficiently — the chemical conditions required for the fuel to power the motor also cause the fuel to decompose wastefully. Both motors currently work in solution, with at

least  $10^{18}$  molecules working in tandem. But the translation of chemical energy into macroscopic motion is likely to require motor components to be constructed in the solid phase, and to be individually controllable.

The story of artificial molecular motors is still in its opening pages, but chemists' attempts to mimic cellular motors reveal how many challenges biological systems have overcome to evolve the machinery that powers movement. The design principles that work are becoming clearer, however, and although the possibility of molecular motors routinely powering artificial devices in the future is still distant, it is now distinct. ■

Jonathan Clayden is at the School of Chemistry, University of Bristol, Bristol BS8 1TS, UK.  
e-mail: j.clayden@bristol.ac.uk

1. Collins, B. S. L., Kistemaker, J. C. M., Otten, E. & Feringa, B. L. *Nature Chem.* <http://dx.doi.org/10.1038/nchem.2543> (2016).
2. Wilson, M. R. *et al. Nature* **534**, 235–240 (2016).
3. Browne, W. R. & Feringa, B. L. *Nature Nanotechnol.* **1**, 25–35 (2006).
4. Kelly, T. R., Sestelo, J. P. & Tellitu, I. J. *Org. Chem.* **63**, 3655–3665 (1998).
5. Davis, A. P. *Angew. Chem. Int. Edn* **37**, 909–910 (1998).
6. Klok, M. *et al. J. Am. Chem. Soc.* **130**, 10484–10485 (2008).

#### PALAEoANTHROPOLOGY

## The dawn of *Homo floresiensis*

**New fossil findings demonstrate that the diminutive hominin *Homo floresiensis* lived on the Indonesian island of Flores at least 700,000 years ago, and may point to its rapid dwarfism from the larger *Homo erectus*. SEE LETTERS P.245 & P.249**

AIDA GÓMEZ-ROBLES

Since the first description of *Homo floresiensis* in 2004 (ref. 1), these little hominins from the Indonesian island of Flores have raised very big questions. Do these skeletal remains represent a new species in the extinct hominin family, or are they modern humans who were pathologically dwarfed, or members of a short-stature population? If they belong to a different species, what was its evolutionary origin? Why was it so different from other hominin species? The most common answer to these questions has been repeated for more than ten years: we need more remains from Flores — especially from different sites and older time periods — to tip the scales. On pages 245 and 249 of this issue, van den Bergh *et al.*<sup>2</sup> and Brumm *et al.*<sup>3</sup> report the finding of those long-awaited remains.

After *H. floresiensis* was described, many palaeoanthropologists embraced the idea of a new and odd-looking hominin species that had

a diminutive brain and body size. Supporters of the pathology hypothesis, however, have been unrelenting in looking for syndromes and conditions that could have been responsible for the unexpectedly small size of these hominins; some suggestions have been published quite recently<sup>4</sup>. The current findings — consisting of a lower-jaw fragment, an indeterminate cranial fragment and some small teeth from at least three different individuals — confirm beyond any reasonable doubt that *H. floresiensis* is a distinct hominin species with deep evolutionary roots that trace back more than 700,000 years.

Van den Bergh and colleagues present detailed analyses of the size and shape of the fossils, found at the Mata Menge site on Flores, including comparisons with remains of other hominin species. They show that such tiny teeth are found only in *Homo sapiens* — whose origin and migration to Asia are substantially later than the age of the new fossils — and in *H. floresiensis*. Brumm and colleagues report

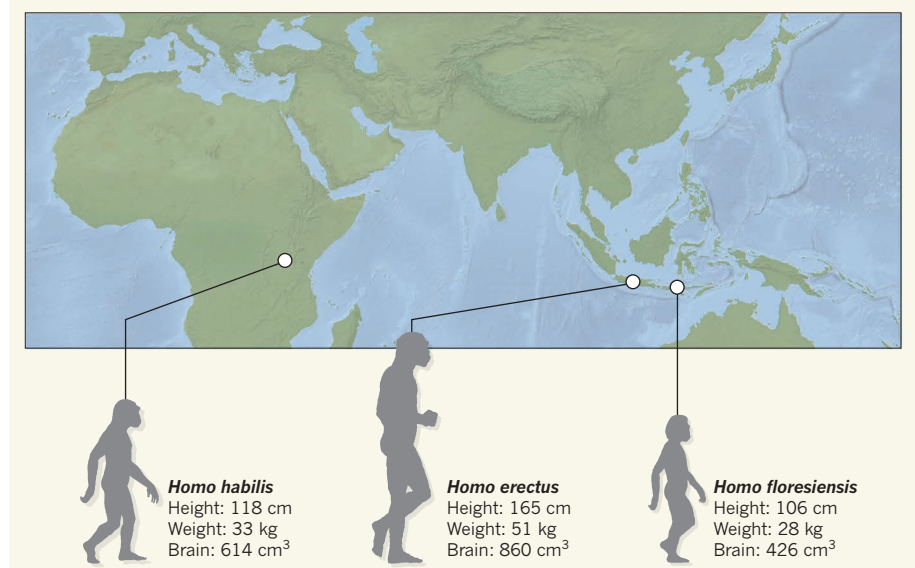
on the open-grassland habitat and stone tools associated with these hominins. They describe these tools as technologically similar to the ones found with the later *H. floresiensis* individuals from the Liang Bua site, and suggest that this points to the behavioural stability of the hominins from Flores over a long period of time. In addition, Brumm *et al.* use a combination of dating techniques to provide evidence that the fossils were deposited around 700,000 years ago, thus confirming the early origin of this species.

Although this confirmation finally ends the debate about the validity of *H. floresiensis* as a species, its evolutionary origins are likely to remain under discussion for much longer. There are two main models (Fig. 1). *H. floresiensis* may have evolved from the larger *Homo erectus* through a process of island dwarfing — an extreme reduction in size due to the absence of predators and to resource scarcity that is typical of island ecosystems. Alternatively, it may be descended from the earlier *Homo habilis*, or even from a small form of hominin from the *Australopithecus* genus. This second model implies that very primitive hominins would have left Africa by 2 million years ago, but there is no fossil or archaeological evidence for such an early dispersal.

Mostly on the basis of the morphology of a lower molar tooth and of general affinities of the jaw fragment, van den Bergh and colleagues claim that the remains from Mata Menge are more closely related to *H. erectus* than to *H. habilis*. The reliability of lower-molar morphology to assess species relationships supports their claim<sup>5</sup>. However, the traits that point to a more primitive ancestor for *H. floresiensis* mostly come from body parts other than the skull<sup>6,7</sup> and cannot be assessed using the Mata Menge sample, which does not include such postcranial remains.

Without further fossil evidence, the discussion between proponents of the two models will continue. Some will think that extreme dwarfing from *H. erectus* is unlikely, especially to the extent of the dramatic brain-size reduction observed in *H. floresiensis*<sup>8</sup>, although empirical data from hippopotamuses suggest that similarly strong brain reduction may occur<sup>9</sup>. Others will argue that a long-distance migration route for *H. habilis*, or an earlier form, from Africa to southeast Asia is even more implausible. For now, it seems that all possible explanations remain outside the comfort zone of classic scenarios of human evolution.

Van den Bergh *et al.* propose that the hominins from Mata Menge might be descended from the hominins that made stone tools at the site of Wolo Sege<sup>10</sup>, also on Flores, which is dated to approximately 1 million years ago. They further hypothesize that large-bodied *H. erectus* hominins are the ones that made these tools. This speculation could be proved wrong if remains from other small hominins



**Figure 1 | Candidates for the ancestry of *Homo floresiensis*.** There are two main models for the evolutionary origin of the hominin species *H. floresiensis*, which inhabited the Indonesian island of Flores and had a particularly small brain and body. One possibility is that *Homo habilis*, or a similar form that also had a relatively small body and brain, may have left Africa by 2 million years ago and reduced in size even further. But there is no evidence for such early hominins outside Africa. Alternatively, *H. floresiensis* may be descended from the later and larger-bodied *Homo erectus*, for which there is evidence on Java around 1 million years ago and earlier. This second model would involve much greater body and brain reduction over a much shorter period of time. (Data on brain and body size are from refs 12–14, and are based on east African specimens for *H. habilis*, on early Indonesian specimens for *H. erectus* and on remains from Liang Bua for *H. floresiensis*.)

were found with the tools in the future, but it raises an interesting question: can the extreme reduction of the brain and body of *H. floresiensis* have evolved over a mere 300,000 years?

Three hundred millennia may not seem a 'short' period of time to many readers. However, no other such dramatic transformation in hominin evolution is known to have occurred over a similarly brief timescale. A quantitative analysis and comparison of evolutionary rates across different hominin species and with *H. floresiensis* would lend formal support to this informal observation. Alongside such quantification, it might be helpful to look at more-distant species. Some mammals show evidence of even stronger degrees of dwarfing over substantially shorter periods of time, and extremely fast rates of size reduction in island environments<sup>11</sup>. In addition, we must not rule out the possibility that the direct ancestors of *H. floresiensis* were not the most typical representatives of their species. Indeed, the strange combination of primitive and derived traits in *H. floresiensis* anatomy could be the result of a pronounced founder effect, which occurs when a new population is established from a small sample that does not reflect the parental population's diversity and most-common traits.

Some scenarios that look mind-blowing from our anthropocentric point of view become underwhelmingly conventional when we expand our horizons. Rapid island dwarfism is not extraordinary in nature,

nor is the founder effect or the long-scale migration of species that lack human-like cognitive abilities. Whatever the actual origin of *H. floresiensis*, we will be much closer to an answer if we look beyond hominins in our search for explanations. ■

**Aida Gómez-Robles** is in the Center for the Advanced Study of Human Paleobiology, Department of Anthropology, The George Washington University, Washington DC 20052, USA.  
e-mail: agomezrobles@gwu.edu

1. Brown, P. *et al.* *Nature* **431**, 1055–1061 (2004).
2. van den Bergh, G. D. *et al.* *Nature* **534**, 245–248 (2016).
3. Brumm, A. *et al.* *Nature* **534**, 249–253 (2016).
4. Henneberg, M., Eckhardt, R. B., Chavanaves, S. & Hsü, K. J. *Proc. Natl Acad. Sci. USA* **111**, 11967–11972 (2014).
5. Gómez-Robles, A., Bermúdez de Castro, J. M., Martínón-Torres, M., Prado-Simón, L. & Arsuaga, J. L. *J. Hum. Evol.* **82**, 34–50 (2015).
6. Jungers, W. L. *et al.* *Nature* **459**, 81–84 (2009).
7. Tocheri, M. W. *et al.* *Science* **317**, 1743–1745 (2007).
8. Martin, R. D., MacLarnon, A. M., Phillips, J. L. & Dobyns, W. B. *Anat. Rec.* **288A**, 1123–1145 (2006).
9. Weston, E. M. & Lister, A. M. *Nature* **459**, 85–88 (2009).
10. Brumm, A. *et al.* *Nature* **464**, 748–752 (2010).
11. Evans, A. R. *et al.* *Proc. Natl Acad. Sci. USA* **109**, 4187–4190 (2012).
12. Kubo, D., Kono, R. T. & Kaifu, Y. *Proc. R. Soc. B* **280**, 20130338 (2013).
13. Grabowski, M., Hatala, K. G., Jungers, W. L. & Richmond, B. G. *J. Hum. Evol.* **85**, 75–93 (2015).
14. Baab, K. L. *J. Anthropol. Sci.* **94**; see [go.nature.com/jdyt9](http://go.nature.com/jdyt9) (2016).



# Accounting for reciprocal host–microbiome interactions in experimental science

Thaddeus S. Stappenbeck<sup>1</sup> & Herbert W. Virgin<sup>1</sup>

**Mammals are defined by their metagenome, a combination of host and microbiome genes. This knowledge presents opportunities to further basic biology with translation to human diseases. However, the now-documented influence of the metagenome on experimental results and the reproducibility of *in vivo* mammalian models present new challenges. Here we provide the scientific basis for calling on all investigators, editors and funding agencies to embrace changes that will enhance reproducible and interpretable experiments by accounting for metagenomic effects. Implementation of new reporting and experimental design principles will improve experimental work, speed discovery and translation, and properly use substantial investments in biomedical research.**

Recent studies highlight that differences in human genetics explain only a fraction of observed human phenotypic variation<sup>1,2</sup>. Another well-recognized source of variation is the microbiome, which is the collection of all host-associated microorganisms. It includes not only the bacterial microbiome but also the virome, archaea, the mycobiome (fungi) and meiofauna (for example, protists and helminths)<sup>3–6</sup>. While the majority of the microbiome is located in the gastrointestinal tract, the skin, oral, and genitourinary microbiome are also colonized by complex sets of organisms<sup>7–9</sup>.

In some cases phenotypes are dominantly controlled by genes in either the microbiome or the host (Fig. 1). However, we increasingly recognize that the effects of host and microbial genes on a phenotype are in many cases interdependent. Such examples include effects of specific microbes (that is, a virus or bacteria) that only occur in defined host genetic backgrounds<sup>10–15</sup>. We and others have referred to this concept as ‘host gene plus microbe’, but a more concise descriptor is that these effects are determined by the metagenome. We herein define the metagenome as the sum of all host genes plus all organism genes of the microbiome. The combinatorial effects of host and microbial genes are herein termed metagenomic effects. Thus, as a driver of phenotypes, one must consider the metagenome, in addition to dominant effects of either host or microbiome genes (Fig. 1).

We now appreciate that nearly all aspects of human physiology, as well as model organisms such as mice, are influenced by the microbiome and metagenome. Though the implications for this work are exciting, it has generated a series of unintended challenges in mammalian experimental biology, especially as related to experimental reproducibility (Box 1). A core issue is a lack of accounting for the microbiome and metagenome, which in turn leads to inconsistent experimental design and interpretation. An underappreciated aspect of the microbiome and metagenome effects is that they are broadly relevant and impact many diverse fields. Below, we address key scientific aspects of this rapidly evolving and exciting area, the relevance and impact of the microbiome and metagenome on physiology and disease as well as specific environmental variables that influence mammalian physiology. We then highlight critical challenges that are created by the influence of the microbiome and metagenome and propose solutions that require the commitment of many diverse stakeholders, including scientists in all areas of investigation, educational institutions, funders of research and journals.

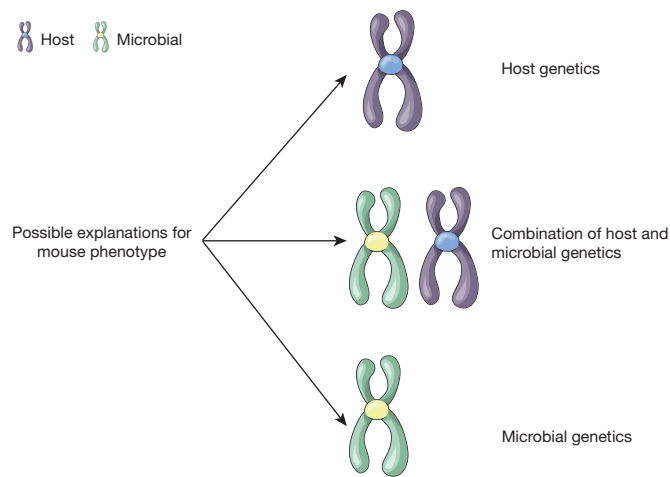
## Metagenome effects central to experimental biology

The recognition that the microbiome and metagenome is a critical factor in human health is not a new concept. Elie Metchnikoff, over 100 years ago, recognized the importance of the microbiome and performed extensive experiments in this area<sup>16</sup>. His thoughts and work foreshadowed the need for new tools to disentangle the complex intestinal microbiome as well as the profound influences of the gut microbiome on systemic organs. The latter concept has recently been demonstrated by the linkage of specific intestinal bacterial metabolites to atherosclerosis<sup>17</sup>.

Important associations between the microbiome and disease have been made. In many cases, associations in mouse models of disease mirror effects observed in humans. These studies support the idea that the microbiome and metagenome have profound local as well as systemic effects, both positive and negative in disease and on physiologic processes related to disease. Examples include inflammatory bowel disease<sup>18–21</sup>, AIDS<sup>22–25</sup>, arthritis<sup>26</sup>, nutrition<sup>27–29</sup>, graft versus host disease<sup>30</sup>, obesity and metabolism<sup>29</sup>, diabetes<sup>31</sup>, haematopoiesis<sup>32</sup>, brain function<sup>33–36</sup>, cancer<sup>37</sup>, bone mass<sup>38</sup> and treatment with immunomodulatory drugs<sup>39</sup>. Thus, alterations of the overall taxonomic complexity, as well as the representation of specific taxa, are associated with a wide range of disease states. Notably these studies show that the microbiome influences systemic organs as well as mucosal organs; no area of physiology is likely to be independent of these influences.

These correlative studies do not demonstrate cause and effect, leading to the development of experimental systems, many in mice, to test mechanistic hypotheses. Improved tools for metagenome analysis, new microbial culture methods and standard studies of host–pathogen and host–commensal interactions have led to the development of novel ideas explaining disease pathogenesis. This has allowed investigators to test for the role of specific microbes or microbial products. For example, microbes and microbial metabolites (i) affect T cell and macrophage differentiation to impact disease states<sup>40–43</sup>; (ii) trigger inflammatory phenotypes in the intestine of genetically susceptible mice<sup>13–15,44</sup>; (iii) alter vascular and crypt structure and development in the intestine<sup>45–47</sup>; (iv) alter anti-viral T cell responses<sup>48</sup>; (v) reactivate herpesvirus from latency<sup>49</sup>; and (vi) affect cardiovascular disease<sup>50</sup>, asthma<sup>51,52</sup>, and infectious disease pathogenesis<sup>6,53</sup>. Further, sequential infection of barrier-raised mice with microbes related to common human infections (herpesviruses, influenza, intestinal

<sup>1</sup>Department of Pathology and Immunology, Washington University School of Medicine, Campus Box 8118, 660 South Euclid Avenue, St Louis, Missouri 63110, USA.



**Figure 1 | The role of the metagenome on determining phenotypes.**

The metagenome is defined as the total host genome (depicted by purple chromosomes) and the associated microbiome genome (depicted by the green chromosomes). It is important to realize that phenotypes can be dominantly driven by either host or microbial genes but also by combinations of genes within the metagenome<sup>13,14</sup>.

helminth) alter gene expression patterns in the blood to partially resemble those observed in adult humans and can alter vaccine responses<sup>54</sup>. Together these studies portend the development of even more extensive data linking specific components of the microbiome to important physiologic mechanisms also involving host genes and improve mouse models as tools for discovery of disease mechanisms and therapeutics.

### Microbiome influence on mouse phenotypes

Mice are amongst the most commonly used model organisms. One common concern is the lack of reproducibility of mouse experiments even within an apparently identical host chromosomal background. In our view this reflects that investigators simply do not consider all of the relevant genes (that is, the metagenome) and their interactions (Fig. 1). The advent of enhanced techniques to genetically modify mice has only accelerated this problem in recent years. Clearly, a variety of environmental factors, in addition to mouse genetic background, such as ambient temperature, water treatment, diet, and light–dark cycles can also impact mouse phenotypes (see below). More recently, many of these same environmental parameters have been linked to alterations of the microbiome. Importantly, these new findings indicate an exquisite interplay between environmental variables, the microbiome, and animal phenotype, creating a veritable ‘witch’s brew’ of concerns about how to interpret what once might have been considered simple experiments.

To clarify the compelling need to confront challenges presented by the emerging recognition of the importance of the microbiome and metagenome, we provide a non-exhaustive list of linkages of the microbiome to variables known to influence experimental outcomes. In each of the cases below, investigators from one facility may not be able to compare their results to those of others if they are unaware of, or cannot duplicate, the conditions under which experiments were performed. This results in apparent lack of reproducibility of findings when in fact, there is not a problem with the experiments *per se*; rather there are unappreciated microbiome variables that influence experimental outcomes.

#### Linkage of the microbiome to host characteristics presumed to be primarily driven by host genetics:

**Example 1.** When phenotypes are passed from one generation to the next it is most commonly assumed that the mode of inheritance is via host chromosomes. The microbiome of mice is largely inherited from the mother (dam), a fact generally attributed to both processes around birth, co-housing and breast feeding<sup>55,56</sup>. Importantly, major phenotypic differences can be inherited (also called vertical transmission) from the

dam in a microbiome-dependent fashion<sup>12</sup>. A diagram representing this is presented in Fig. 2 showing that the levels of IgA in faeces can be transmitted vertically (and thus be ‘heritable’) through the microbiome of the dam, and can also be transmitted horizontally from one mouse to another when IgA ‘high’ and IgA ‘low’ mice are co-housed. This phenomenon also occurs with transmission of segmented filamentous bacteria within mouse facilities including vendors<sup>57</sup>.

**Example 2.** Host strain background controls many experimental phenotypes and also plays a role in the microbiome composition<sup>58</sup>. Since the microbiome can control host phenotypes, findings attributed to genetic background may instead be due to effects of the microbiome.

**Example 3.** The gender of mice used for experiments can impact both phenotypes and microbiome. For example, the sex of mice alters diabetes and autoimmune phenotypes associated with differences in the microbiome<sup>59,60</sup>.

**Example 4.** Many experimental phenotypes are related to the age of mice. Age can influence the microbiome<sup>61</sup>, suggesting that some of the effects of ageing may be in part due to changes in the microbiome. Attention should also be paid in the first two months of life, during which time the intestine undergoes rapid changes in development.

**Example 5.** The microbiome undergoes profound circadian changes, and can control epithelial and immune cell homeostasis and metabolism<sup>62,63</sup>. The light–dark cycle of facilities and time of day of experiments and sample collection are often not standardized or reported in manuscripts.

Linkage of the microbiome to environmental variables known to influence mouse biology:

**Example 1.** Intercurrent infection of mice with known pathogens such as, to name but a few, *Helicobacter* species, murine norovirus, mouse rotavirus (epizootic diarrhoea of infant mice), pneumocystis and many others can alter results of experiments<sup>14,15,64</sup>. Testing for such pathogens is commonly unreported in publications. Variations in these infections between mice in different cages and in different facilities have profound biologic implications for reproducibility of results between different facilities and institutions. As an example, persistent norovirus infection of germ-free mice can emulate some of the immune-related developmental effects normally attributed to the commensal intestinal bacteria<sup>44</sup>, and bacteria are critical to the inflammatory effects induced by norovirus infection in mice carrying mutations in genes associated with risk for human inflammatory bowel disease<sup>14,15</sup>.

**Example 2.** It is not often appreciated by investigators outside of metabolism research that prevention of cold stress in mice requires an ambient temperature of 26–29 °C. Most facilities maintain temperatures of 20–24 °C for the comfort of the investigators<sup>65</sup>. To compensate for cold facilities, mice will titrate nesting material resulting in potential changes in mouse phenotype based on bedding availability<sup>66</sup>. Cold stress of mice can profoundly impact mouse physiology including elevating heart rate and alterations of basal metabolism driven in part driven by excessive glucocorticoid production. Chronic glucocorticoid stimulation is immunosuppressive which in turn has a profound impact on many types of experiments including responses to pathogens and the development of immunity. Chronic glucocorticoid treatment also changes the intestinal microbiome<sup>67</sup>, indicating that effects of cold stress might be on, or through, changes in the microbiome. This issue has been addressed experimentally by intentionally subjecting mice to extreme cold stress. This environmental perturbation alters the microbiome in mice, and this alteration can be transmitted to mice not exposed to cold temperatures. Importantly, the microbiome changes contribute to phenotypes<sup>68</sup>. Another conclusion from this body of work is that the common practice of transporting mice from mouse facilities to the laboratory in containers that do not include bedding or that travel through cold (or hot) environments may alter results.

**Example 3.** Water treatment and sterilization can alter both phenotypes and the microbiome. Some facilities rely on sterilization by autoclaving, while others rely on acidification. The pH of the water in some instances is ~2, which can alter both mouse phenotypes and the microbiome<sup>69,70</sup>.



## BOX 1

## Challenges created by metagenomic influences on mouse phenotypes

## Overall scope

- Failure to recognize the proven impact of the microbiome and metagenome on the acquisition and interpretation of experimental data in research planning, grants and publications.
- Pervading view in many fields that issues related to the microbiome and metagenome affect only mucosal biology and immunity. This view is not supported by recent findings that show the marked global impact of the microbiome on many areas of host physiology.
- The effects of the microbiome and metagenome are biologically, conceptually and bioinformatically complex, making them challenging to quantify and analyse even if one accepts the need to account for them.

## Experimental design

- Lack of rigorous accounting for environmental factors that influence the microbiome and metagenome in experimental design and interpretation. This will hamper collaboration both within, and increasingly between, countries and continents, thereby limiting development of global science by infusing lack of reproducibility of apparently similar experiments.
- Incomplete recognition of the fact that the metagenome and microbiome may affect experimental results via both vertical and horizontal transmission<sup>12</sup>.
- An erroneous (and common) assumption that descriptive studies involving sequencing and analysis of the intestinal bacterial microbiome is sufficient to control for the effects of the metagenome.
- Lack of understanding of substantial variations between the mouse facilities and resources available to investigators at different institutions to discern them.

## Experimental analysis

- Lack of facile centralized databases for mouse phenotypes and microbiome sequence data sets that are comparable to those used in structural biology or gene expression analysis to foster proper experimental design, analysis and comparison between data sets.
- Severely limited access of investigators to cutting edge computational tools and expertise needed for analysis of the microbiome. These resources most often cannot be generated by individual research groups.
- Limits of single investigator 'bandwidth'. The expertise of individual investigators is typically restricted to one or a few of the components of the microbiome (for example, the virome, bacterial microbiome or mycobiome).

## Administrative and organizational

- Failure to recognize and address these challenges cannot be attributed to a single group of stakeholders.
- Failure to recognize that solutions to these challenges cannot be implemented by a single group of stakeholders.
- Fully addressing the impact of the microbiome and metagenome on experimental findings is expensive and resource intensive.
- Delays in productivity attendant upon full evaluation of the microbiome and metagenome in experimental systems may influence promotions and grant competitiveness, thereby inhibiting professional development and training.
- Lack of consensus across different stakeholder groups of how to deal with these challenges.

In this work we highlight the importance of the microbiome in altering mouse phenotypes. Importantly, these effects are not exclusively local (that is, reserved to the intestine) but instead can have profound effects on diseases that occur in distant organs. These facts present many challenges (listed here) for scientific enterprise. What is clear is that there is no simple, single-stakeholder solution. In the latter part of this Perspective, we provide a series of experimental and policy recommendations to begin to overcome these obstacles. We feel that the enormous potential for both scientific understanding and improving the human condition of defining mechanisms responsible for the impact of the microbiome and metagenome on physiology and diseases will not be fully realized unless these challenges are addressed.

**Example 4.** Diet has profound effects that can outweigh host genetics<sup>71</sup>. The composition and handling of mouse chow, with alterations in composition of fat or simple sugars alters both the microbiome and mouse phenotypes<sup>71</sup>. Other additions to the diet, such as haem (to mimic red meat), can impact inflammatory phenotypes and the microbiome<sup>72</sup>.

**Example 5.** Transport of mice from large dedicated breeding/repositories facilities is frequently used, particularly to obtain wild-type controls. In addition to probably mismatched microbiomes (because mice are from different facilities), the travel-related stress to animals (for example, temperatures in trucks and planes, handling, time of travel) cannot be controlled, and it is not known how long mice take to adjust to the new environmental and microbial milieu of a facility. This complicates and may invalidate comparisons of mutant mice bred in a research facility with wild-type control mice from an outside vendor.

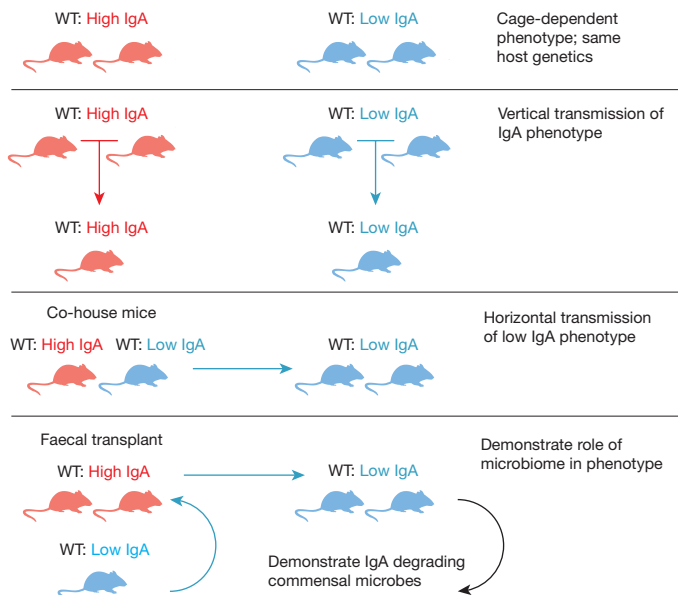
## The case for littermate controls

Well-established methods exist to control for host, microbiome and metagenomic effects (Fig. 1). These methods have an increasingly appreciated scientific basis, but unfortunately they are not always used. As an example, consider the common practice of comparing wild-type mice to mice with a mutation when the wild-type and mutant mice are bred separately or one group is purchased from an outside facility (Fig. 3). The results of such comparisons are often reported in the literature as

conclusive evidence that a mouse genomic mutation caused the phenotype. However, this is not a valid conclusion. When mice are bred separately, the microbiome can confer apparently heritable phenotypes<sup>12</sup> and phenotypes can be transferred between a mutant and a wild-type mouse by co-housing<sup>12,73–75</sup>. To complicate this scenario, both viruses and bacteria can confer phenotypes in the presence of specific host gene mutations through metagenomic effects<sup>13–15,44</sup>.

A well-recognized solution to this problem is to use littermate controls. For autosomal genes, breeding heterozygous parents controls for the effects of the microbiome (Fig. 3). Breeding heterozygous parents yields all of the groups required for phenotypic comparison when combined with tracking of which mice come from which dam to control for phenotype inheritance via the maternal microbiome<sup>12</sup>. In the analysis of adult mice, use of littermate controls addresses both maternal and early life effects of the microbiome which can imprint the immune and metabolic systems<sup>76,77</sup>. It is the gold standard to determine the dominance of host versus microbiome genes in conferring a phenotype.

Importantly, littermate controls are required for studies of host genes that impact organs distant from highly colonized organs such as the intestine. This generalizes the importance of the microbiome in experimental design. Metabolites produced in colonized organs can directly influence phenotypes at distant sites including lung<sup>52,78</sup>, pancreas<sup>79</sup>, and brain<sup>33,35</sup>. For example, the intestinal bacterial microbiome plays a key

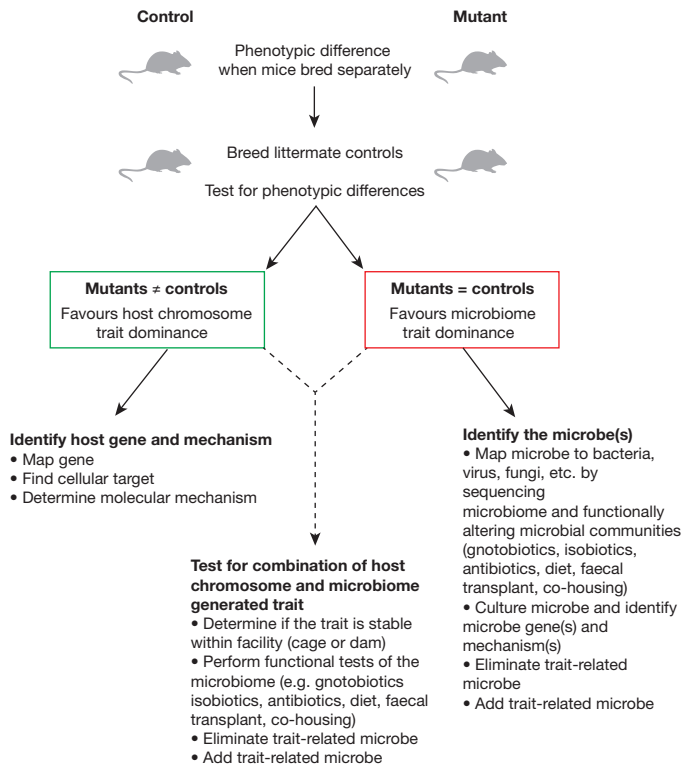


**Figure 2 | Determining the role of the microbiome within genetically equivalent but phenotypically different mice (as in ref. 12).** A common occurrence is that mice of the same genotype within a given colony can show differences in phenotype. In this example, levels of faecal IgA show variation in wild-type (WT) mice that were cage dependent: 'High IgA' denotes detectable levels of faecal IgA (as determined by an ELISA), while 'Low IgA' denotes levels at the limit of detection of the assay (top panel). Here, this phenotype is vertically transmitted. IgA-high mice produce progeny that are IgA high and IgA-low mice produce progeny that are IgA low (middle panel). Two experiments using horizontal transfer methods show dominance of the IgA-low phenotype. First, co-housed IgA-high and -low breeding pairs produce mice that are IgA low. Second, faecal transfer from IgA-low to IgA-high mice converts IgA-high mice to IgA-low (bottom panel). This experiment then generated the material required to determine that IgA was degraded by intestinal commensal microbes. These experiments show that the IgA-low phenotype is dominant and driven by the microbiome.

role in resistance to pulmonary influenza virus infection<sup>80,81</sup>. In addition, mobile cells such as those in the immune system can modify phenotypes in distant non-colonized organs<sup>82</sup>. Thus, experimental design in all areas of experimental biology requires consideration of microbiome effects and thus the utilization of littermate controls. These considerations conclusively demonstrate that a simple comparison of mutant and control mice bred independently or purchased from a facility cannot be interpreted as definitive evidence for the role for a host gene. Comparisons between mutant and littermate control mice can, however, identify whether the host or microbiome is dominant in controlling the trait under study.

Importantly, using littermate controls does not rule out a role for metagenomic effects. Specifically, it may still be the microbiome that dominantly modulates a trait that is also dependent on a host gene (or vice versa; Figs 1 and 3). To address this possibility, a next step in cases with a dominant host genetic contribution is to define the potential role of the microbiome by evaluating the phenotype using additional dams, assessing variation between cages<sup>55</sup>, using faecal transplantation or co-housing, and by analysing mice re-derived into a separate facility (Fig. 4). These studies can demonstrate the impact of a combination of host and microbial effects that must be interrogated to define the mechanisms that underlie experimental observations (Fig. 3).

Many studies demonstrate that phenotypes can be transferred between mice by either co-housing or faecal transplantation<sup>11–14,73,75</sup>. Investigators should be aware that using this valuable approach in adult mice can address the effects of the microbiome independent of effects on development. In our opinion, these experiments supplement, but do not replace, the use of littermate controls (Fig. 3). In addition to their value for determining whether a phenotype is influenced by the microbiome after



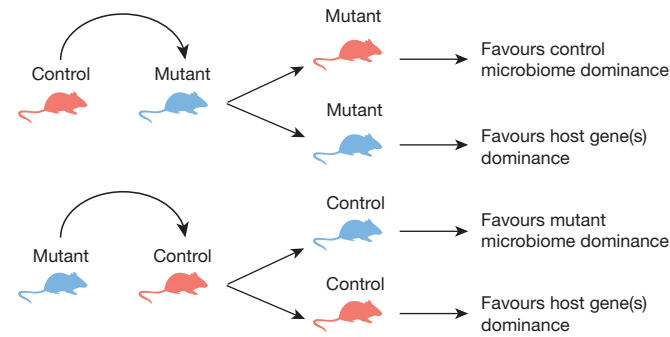
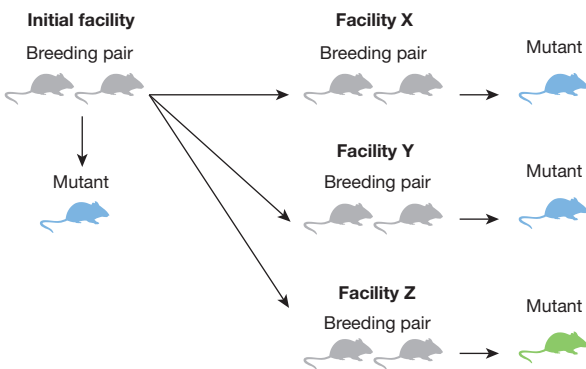
**Figure 3 | Use of littermate controls as a gold standard to control for the effects of the microbiome and metagenome.** When control and mutant mice are bred independently, a difference in phenotype cannot be simply attributed to the difference in host gene(s). This applies to all physiologic contexts. Experiments that use littermate controls must be performed to define whether the phenotype is dominantly controlled by host gene(s), microbial gene(s) or a combination of both types of genes. If the results of such experiments show the mutant mouse phenotype is still distinct from controls, then the trait is probably dominantly controlled by host chromosomal genes. If the mutant mouse phenotype resembles the littermate controls regardless of host genotype, then the trait is probably dominantly controlled by the microbiome. It is important to note that these conclusions are not definitive; the possibility will still exist that a combination of host and microbiome (metagenomic) influences may drive a given trait. Additional experiments must be performed to test for this possibility (see additionally Fig. 4).

development of the animal, these studies can provide the directionality of the influence (Fig. 4a). The mechanism of microbial dominance can be through effects of individual microbes, changes in community structure or combinations of these with host genes<sup>13,14,73,83</sup>.

Use of isobiotic mice<sup>84,85</sup>, those with exposure to only a defined set of organisms, is also an attractive experimental system for analysis of microbiome effects. Germ-free animals reconstituted with defined bacteria, human microbiome samples, or the microbiome of conventionally raised mice also have a key role in defining the role of the microbiome in physiology, although the effects of the microbiome on normal development are not addressed with such experiments<sup>19,29,86</sup>. Such experimental systems have been important for establishing the role of host microbial communities in modulation and host phenotype.

Lastly, to most conclusively address the role for the metagenome, certain key experiments should be performed in multiple animal facilities in order to draw firm conclusions about the generality of a role of host and/or microbiome genes in a phenotype. Such experiments are particularly critical for new and foundational areas of biology. Just as the clinical world relies on multi-centre trials as its gold standard for treatment efficacy, multi-centre vivarium studies, sometimes as a follow-up of initial observations, would solidify the conclusions regarding the generalizable role of the microbiome in a given phenotype (Fig. 4b). Such an approach has the potential to eliminate the confusion regarding phenotypes erroneously



**a Faecal transplant between genotypes to test for dominance****b Facility distribution of breeding pairs to test initial phenotype**

**Figure 4 | Additional experiments to test for the relative dominance of host chromosomal or microbiome traits.** **a**, In addition to testing for littermate controls (see Fig. 3), faecal transplant between control and mutant mice can be performed. If the host gene is dominant, the phenotype will not be transferred horizontally or by microbiome transplantation. If a microbe (or microbes) is (are) dominant, then the phenotype will be transferred from the mice of one genotype to the other genotype. We note that co-housing is also a viable method and in the event that co-housing transfers a phenotype but faecal transplant does not, it is possible that the phenotype is due to a microbiome component (for example, enveloped virus, anaerobic bacteria) that does not survive the transplantation procedure but may efficiently spread between live animals. Mouse genotype is indicated by text (control or mutant), phenotype is indicated by colour: red, normal; blue, altered. **b**, A further test to evaluate the generality of the effects of the metagenome on a phenotype is to transfer mice to additional mouse facilities and determine the phenotype in those facilities. This may confirm the phenotype in the initial facility (blue box) or not (green box). If not, this is an opportunity to define why there is an altered phenotype by comparing the environment between different facilities in a controlled manner.

linked to genotypes that has befuddled many fields, wasting time and critical resources as experiments are performed to resolve apparent controversies. For robust mouse phenotypes that define a specific disease model, such meta-analysis across multiple facilities would be ground-breaking and diminish confusion surrounding the variability of certain models. There is much folklore in the biology community with regard to reasons for lack of reproducibility, but little scientific action. There are several examples of low-hanging fruit that should be addressed sooner than later. For example, the facility-dependent variation in the onset of diabetes in non-obese diabetic mice is a major impediment for the field<sup>87,88</sup>. In addition to investigator collaboration, funding agencies will have to recognize the utility of this endeavour by providing funding, as will high-impact journals by prioritizing studies that meet this demanding standard.

### Role of microbiome sequencing

Based on rapid advances in sequencing technologies, improvement in databases containing sequences from microorganisms and development

of the bioinformatic capacity to map sequences to the genomes of specific types of microorganisms, it has become increasingly clear that mammals are composite organisms containing genes not only from host cell chromosomes, but genes from many microbes<sup>3,4,6,7,9,18,89–91</sup>. In our opinion, however, simply reporting the ‘sequence’ of the microbiome does not solve the problem of incorporating effects of the microbiome into analysis of experimental findings. This is because the ‘sequenome’, meaning the total sequences that are obtained from sequencing random nucleic acid fragments from the microbiome (shotgun libraries), does not completely describe the metagenome. There are multiple reasons for this mismatch. First, not all elements of the microbiome (that is, archaeal, bacterial, viral, fungal, etc.) can be analysed by the same sample preparations. This is a challenge for limited samples as well as for bioinformatic comparisons of data sets derived from different sample preparations. Second, shotgun sequence libraries contain many sequences that cannot be annotated to any organism in current databases. This genetic ‘dark matter’ probably contains sequences from as-yet-unidentified, but physiologically important, components of the microbiome. Inability to characterize the entire sequenome is a key limitation to current studies of the metagenome. Third, functional microbes may be of such relative low abundance that they may not be accounted for by sequencing<sup>92</sup>. Lastly, the functional impact of metabolites or particles such as outer membrane vesicles will not be detected by sequencing<sup>93</sup>.

Analysis of the bacterial microbiome has advanced more rapidly than analysis of other organisms within microbiome. This has occurred in part because of robust databases for bacterial 16S rRNA genes or programs that link shotgun sequences to specific bacterial taxa (that is, MetaPhlAn<sup>94</sup>) allowing the investigator to correlate sequence with the presence of specific bacteria (often down to the level of genus or species). However, because analysis of the bacterial microbiome has developed more rapidly than analysis of the metagenome, there has been an unfortunate tendency in the literature and in scientific presentations (and thought) to use the term ‘microbiome’ to refer to bacteria only, ignoring other types of organisms. In addition, more attention has been paid to microbiome members with DNA genomes than those with RNA genomes due to the use of libraries made without reverse transcription. There are many members of the microbiome with RNA genomes; for example many enteric RNA viruses are detected in apparently healthy mammals<sup>3,6</sup>.

The number of organisms, and therefore the number of genes, constituting the microbiome is enormous. In the intestine alone, it is estimated that there are about 100-trillion prokaryotic cells at densities of up to  $10^{11}$ – $10^{12}$  cells  $\text{ml}^{-1}$  in the colon (fewer more proximally)<sup>95</sup>. Bacteriophages that infect these prokaryotes may be in the range of ~tenfold more abundant than their respective host cells<sup>3</sup>. Further, humans carry an average of at least ten permanent chronic eukaryotic viral infections that substantially influence immunity and host gene expression<sup>6,96</sup>, and faecal samples from healthy children contain a range of eukaryotic viruses including members of genera that contain pathogens<sup>3</sup>. Mammalian chromosomes contain multiple genetic elements related to retroviruses, and recombination amongst these can generate new infectious retroviruses in mice, indicating that our own genomes may be the source of novel elements of the virome<sup>97</sup>. Members of the virome may also be living within other eukaryotic organisms, such as parasites, that infect the host<sup>98</sup>. The microbiome also contains a fungal mycobiome<sup>99</sup> and meifauna such as protozoa and helminthic worms<sup>3,100,101</sup>. Thus, defining two mice as having equivalent microbiomes by shotgun sequencing is currently not possible, and the field is too immature to mandate specific standards for microbiome analysis due to the rapid evolution of this discipline.

### Should we standardize the microbiome?

It is clearly challenging to choose a set of even bacteria for a given set of experiments (for example, using isobiotic mice). Further, this approach will miss the critical effects of other components of the microbiome (viruses, etc.). In addition, it is important not to assume that the different components of the microbiome act independently; instead trans-kingdom interactions between components of the microbiome can dramatically

affect the biology of the host<sup>102</sup>. Bacteriophages participate in horizontal gene transfer between prokaryotes and participate in predator–prey relationship creating a dynamic community structure comprised of the bacterial microbiome and the bacteriophage component of the virome<sup>19,103–105</sup>. The induction of replicative bacteriophages which can infect other bacteria can be induced by environmental factors such as nitric oxide, antibiotics, and nutrient availability<sup>106</sup>. Thus, the environment can control the bacteria microbiome via effects on the virome.

In turn, the bacterial microbiome and individual bacterial products can regulate the ability of eukaryotic viruses to establish both acute and persistent infection in mice<sup>102,107–111</sup>. Chronic systemic viral infections control the level of innate and adaptive immunity to bacteria, parasites, other viruses, tumours, and regulate autoimmunity<sup>6,10,112</sup>. Vaccine responses to viruses are controlled by the microbiome<sup>54,113–115</sup>.

The intestinal meiofauna is complex and present in asymptomatic individuals<sup>3,99,116,117</sup>, and may alter inflammatory bowel disease, multiple sclerosis, rheumatoid arthritis, type 1 diabetes and asthma<sup>118</sup>. As an example, sequencing the mycobiome in faecal samples from apparently healthy individuals detected fungi in every sample tested and over 50 fungal genera<sup>119</sup>.

This level of taxonomic complexity, and the existence of trans-kingdom interactions within the metagenome that have important biological effects, means that defined microbiomes cannot yet be a standard for all experiments if we expect biological experimentation to encompass all mechanisms that operate in human populations. Focusing on a single ‘standard’ microbiome will prevent analysis of major variations in biology that occur when the microbiome varies in the same chromosomal genetic background, and thus important opportunities will be missed to discover new interactions of the microbiome with the host.

### Addressing challenges of metagenomic effects

Given all of the above, we are confronted, as investigators and recipients of research support, with the problem of how to deal with these complex but now proven-pertinent considerations. We feel that significant problems contributing to lack of experimental reproducibility, proper interpretation of experimental findings and our failure to realize the full potential of studies of the microbiome and metagenome to generate new paradigms in biology cannot be simply solved by dictating and enforcing ‘favoured’ experimental procedures for design of mouse experiments. While we do strongly favour certain methods based on scientific evidence (see description of littermate controls and other valuable approaches above), proposing ‘rules’ meant to rigidly apply to all studies would be unproductive. Instead, we feel strongly that a specific scientific enterprise (such as the use of mouse models), if given the proper tools and framework, can move towards reproducibility without compromising new discoveries or regimenting science.

We have evaluated how other scientific communities that rely on comparability and integrity of data across institutions and investigators deal with these issues. We are particularly impressed with the fast evolution of the expression microarray field. In the early days of this field (the late 1990s), there was much concern that data from experiments could not readily be compared between laboratories or platforms and there were substantial concerns about the overall quality and utility of the data. Instead of dictating a single favoured platform, the field embraced the complexity of technical development of platforms by simply reporting how things were done. There were two key proactive developments. First was the insistence that all parameters of an experiment be reported (the consensus became reporting minimum information about a microarray experiment, MIAME). This greatly facilitated comparison of platforms and identified trouble spots. This approach was quite effective at streamlining procedures and facilitating cross comparisons across platforms. Second, microarray data was required to be deposited and made freely available upon publication. Journal editors in particular have been important enforcers of such depositions.

Thus, we propose to mandate the creation of reporting systems for all experimental models that will facilitate cross-comparison of experimental

design and data obtained from mouse experiments. While we suggest a set of parameters to be reported (Box 2), we recognize that this list is probably incomplete and will need refinement by the research community. We feel that transparent sharing of information is the missing piece needed to solve this set of problems, and that participation in such a system is an ethical responsibility for investigators that utilize public and private funds. Some microbiome researchers have already embraced this type of transparency and steps are being taken to provide frameworks for reporting of key variables in mouse experiments<sup>120</sup>. Consensus conferences would be the logical next steps to produce a concrete list of information that is required to be shared by investigators (see below).

### Facilitating mouse model reproducibility

The solution of the metagenomics ‘problem’ must be addressed simultaneously from multiple sources and through the actions of different stakeholder groups. This will require recognition of this issue at many levels as well as the motivation to take appropriate action. We delineate the responsibilities of various entities in improving experimental reproducibility and enhancing the value of investments in research below:

(1) *Individual investigators*. It is incumbent on the individual investigator to acknowledge and control for the metagenome in all mouse experiments in all areas of investigation for reasons described in this review. For labs that use mouse models, this level of understanding must be effectively communicated to all members of a laboratory including technicians, graduate students and postdocs. We recommend that littermate controls are the minimum gold standard for experimentation for non-gnotobiotic facilities. We recognize that other methods may be used by investigators and appropriate interpretations as to the role of the microbiome and metagenome must be made based on the method (Figs 3 and 4). Details of mouse experimental parameters must be clearly documented (for example, Box 2); this will allow the community to judge the value of reported findings and attempt to replicate them. If sequencing of the microbiome is undertaken, the methods for generating these data and the primary sequence files must be made freely available. Reported information will provide investigators the ability to evaluate the credibility of the results and place them in a proper context in the field.

(2) *Government and business sectors*. We must define MIAME-like criteria for mouse experiments. Many of the features are touched on in this review, although we recognize that this is not an all-inclusive set of considerations and reflects our own opinions rather than an expert panel or community consensus. Thus input and consensus across disciplines will be required to define required reporting elements. We propose that a series of consensus conferences, to include the key stakeholders, be established to generate and update these MIAME-like criteria. Further, a fully funded commission or committee should be created to monitor the science, study the specific effects of policies put in place, adapt standards for reporting to the literature, and to identify and remediate unanticipated negative consequences of well-meaning policies that will certainly arise.

An essential element of a successful plan will be the development and maintenance of data repositories to allow investigators and others to readily search for mouse phenotypes with a given host background and compare associated experimental details for each experiment. Standardized methods of reporting such information will facilitate and push fields in the direction of best practices that facilitate reproducibility. Such a database will be transformative for many fields as it will finally allow for facile cross comparison of experiments within a defined host genotype. Providing links to other OMICS data collection (that is, transcription microarrays, RNA-seq, microbiome analysis) would increase the power of this approach. Only government agencies, perhaps in combination with industry, are well positioned to foster, fund, and maintain such databases.

(3) *Journal editors and reviewers*. In order to publish, investigators must provide data critical for interpretation and establishing the reproducibility of results. Journals must enforce reporting of key experimental parameters and deposition of data in central databases for mouse phenotypes. Editors and journal policies must provide the ‘teeth’ for this endeavour. Without the commitment of the editorial process, this idea will die, and



## BOX 2

## Example of types of data to be provided for submitted manuscripts with mouse experiments

### Host genetics

- Specify strain using JAX or other commercial vendor nomenclature.
- Original source for purchased or shared mice used for breeders used to create colony.
- For mixed background, include data defining strain percentage (microsatellite analysis, number of markers).
- Define the method used to create the mutation (for example, homologous recombination in embryonic stem cells, transposon mutagenesis, chemical mutagenesis, Cas/CRISPR systems). Show data validating the altered allele.

### Experimental methods within mouse facility

- Source of experimental and control mice (for example, bred in facility, purchased from specified vendor; for latter interval from arrival in facility to experiment).
- Control for microbiome effect (for example, littermates, multiple dams, co-housing, faecal transplant, gnotobiotic).
- Breeding scheme to generate experimental and control mice.
- Number of breeding pairs used to generate progeny for analysis.
- Number and gender of mice analysed per experiment.
- Number of experiments performed.
- Antibiotic exposure (type and duration) of breeders and progeny.

### Husbandry details

- pH of drinking water.
- Diet source (vendor, nutrient composition), storage (temperature, duration) and treatment (irradiation, autoclave).
- Caging type (for example, ventilated, metabolic).
- Bedding amount per cage and type.
- Frequency and protocol for cage changing.
- Light–dark cycle of room.
- Temperature of room (include range).
- Pathogen screening (organisms tested for, methods, source of analysis in house versus commercial vendor).

### Microbiome analysis

- Methods of sample collection, library preparation.
- Analytical pipeline including version and database dates.
- Methods of statistical analysis.
- Specify method used if corrections for multiple comparisons were performed.

the reproducibility and quality of science will continue to be less than optimal. The pressure to provide this information in order to publish is a very powerful motivational force. Furthermore, in the review process, editors will need to give precedence for publication to studies that address effects of the microbiome and metagenome over studies that fail to do so. The end product will make experimental conditions transparent and comparable which will have a profound effect on the value of the scientific enterprise to the public.

(4) *Academic institutions.* The leadership of academic institutions (for example, presidents governing boards, chancellors, provosts, deans, department and division heads) must develop directed plans to create the infrastructure and faculty expertise that is required to properly perform animal experimentation in a metagenomic world. This will require money. It is critical that individual investigators using mouse models have the expertise available within their institution or through consortia to address issues of the metagenome. This includes access to computational programs, methods to evaluate the microbiome as well as experts in computational biology and various disciplines of microbiology (that is, bacteriologists, virologists).

(5) *Granting institutions, study section heads and reviewers.* Grant proposals must document practices and controls for mouse experiments. Adherence to defined community-designed and -supported best practices will need to be promised and acted upon. Experimental procedures that account for the microbiome and metagenome must be described. Institutional resources available to the investigator must be documented so that it is clear how these issues will be addressed. Lastly, plans to deposit information about experimental parameters and data deposition must be described. These details should be required for all grant proposals that use mouse models.

## Conclusion

We expect that some of the opinions expressed here will be controversial. However, while hard to implement and expensive, we believe that the data are now conclusive that such efforts are required to optimize experimental biology in the coming decades. The first step will be recognition and acceptance of the importance of the effects of the microbiome and metagenome. This can only come if investigators have access to the information for other experiments in the literature that is required for interpretation and reproduction of findings. In the end, the microbiome and metagenome are coming into their own as a field driving innovation in experimental biology. We believe that the complexity of the metagenome should be embraced for its potential to change how we view ourselves and fellow mammals as organisms, and for the potential to truly understand who we are and how to better our lives. After all, we are composite organisms containing far more genes than are encoded in our own chromosomes.

Received 25 November 2015; accepted 26 April 2016.

1. Brodin, P. *et al.* Variation in the human immune system is largely driven by non-heritable influences. *Cell* **160**, 37–47 (2015).
2. Lee, M. N. *et al.* Common genetic variants modulate pathogen-sensing responses in human dendritic cells. *Science* **343**, 1246980 (2014).
3. Norman, J. M., Handley, S. A. & Virgin, H. W. Kingdom-agnostic metagenomics and the importance of complete characterization of enteric microbial communities. *Gastroenterology* **146**, 1459–1469 (2014).
4. Moon, C. & Stappenbeck, T. S. Viral interactions with the host and microbiota in the intestine. *Curr. Opin. Immunol.* **24**, 405–410 (2012).
5. Hooper, L. V., Littman, D. R. & Macpherson, A. J. Interactions between the microbiota and the immune system. *Science* **336**, 1268–1273 (2012).
6. Virgin, H. W. The virome in mammalian physiology and disease. *Cell* **157**, 142–150 (2014).

7. Beck, J. M. *et al.* Multicenter comparison of lung and oral microbiomes of HIV-infected and HIV-uninfected individuals. *Am. J. Respir. Crit. Care Med.* **192**, 1335–1344 (2015).
8. Anahtar, M. N. *et al.* Cervicovaginal bacteria are a major modulator of host inflammatory responses in the female genital tract. *Immunity* **42**, 965–976 (2015).
9. Oh, J. *et al.* Biogeography and individuality shape function in the human skin metagenome. *Nature* **514**, 59–64 (2014).
10. MacDuff, D. A. *et al.* Phenotypic complementation of genetic immunodeficiency by chronic herpesvirus infection. *eLife* **4**, (2015).
11. Virgin, H. W. & Todd, J. A. Metagenomics and personalized medicine. *Cell* **147**, 44–56 (2011).
12. Moon, C. *et al.* Vertically transmitted faecal IgA levels determine extra-chromosomal phenotypic variation. *Nature* **521**, 90–93 (2015).  
**Demonstrates that indigenous intestinal bacteria dominantly influence phenotypes through degradation of fecal IgA and emphasizes the need for control of the metagenome and microbiome in mouse experiments.**
13. Bloom, S. M. *et al.* Commensal *Bacteroides* species induce colitis in host-genotype-specific fashion in a mouse model of inflammatory bowel disease. *Cell Host Microbe* **9**, 390–403 (2011).
14. Cadwell, K. *et al.* Virus-plus-susceptibility gene interaction determines Crohn's disease gene Atg16L1 phenotypes in intestine. *Cell* **141**, 1135–1145 (2010).  
**With ref. 13, experimental proof of concept that one cannot only separately consider the effects of host genes or microbiome genes on mouse phenotypes; one must consider the entire metagenome including bacteria and viruses.**
15. Basic, M. *et al.* Norovirus triggered microbiota-driven mucosal inflammation in interleukin 10-deficient mice. *Inflamm. Bowel Dis.* **20**, 431–443 (2014).
16. Metchnikoff, O. *Life of Elie Metchnikoff 1845–1916*. (Houghton Mifflin Company, 1921).
17. Wang, Z. *et al.* Non-lethal inhibition of gut microbial trimethylamine production for the treatment of atherosclerosis. *Cell* **163**, 1585–1595 (2015).  
**Demonstrates the concept that intestinal metabolites can have dramatic effects on distant organ systems (here blood vessels) and emphasize that the effect of the metagenome and microbiome must be considered in all mouse experiments.**
18. Gevers, D. *et al.* The treatment-naïve microbiome in new-onset Crohn's disease. *Cell Host Microbe* **15**, 382–392 (2014).
19. Reyes, A., Wu, M., McNulty, N. P., Rohwer, F. L. & Gordon, J. I. Gnotobiotic mouse model of phage-bacterial host dynamics in the human gut. *Proc. Natl Acad. Sci. USA* **110**, 20236–20241 (2013).
20. Norman, J. M. *et al.* Disease-specific alterations in the enteric virome in inflammatory bowel disease. *Cell* **160**, 447–460 (2015).
21. Haberman, Y. *et al.* Pediatric Crohn disease patients exhibit specific ileal transcriptome and microbiome signature. *J. Clin. Invest.* **124**, 3617–3633 (2014).
22. Handley, S. A. *et al.* Pathogenic simian immunodeficiency virus infection is associated with expansion of the enteric virome. *Cell* **151**, 253–266 (2012).
23. Mutlu, E. A. *et al.* A compositional look at the human gastrointestinal microbiome and immune activation parameters in HIV infected subjects. *PLoS Pathog.* **10**, e1003829 (2014).
24. Handley, S. A. *et al.* SIV infection-mediated changes in gastrointestinal bacterial microbiome and virome are associated with immunodeficiency and prevented by vaccination. *Cell Host Microbe* **19**, 323–335 (2016).
25. Monaco, C. L. *et al.* Altered virome and bacterial microbiome in human immunodeficiency virus-associated acquired immunodeficiency syndrome. *Cell Host Microbe* **19**, 311–322 (2016).
26. Scher, J. U. *et al.* Expansion of intestinal *Prevotella copri* correlates with enhanced susceptibility to arthritis. *eLife* **2**, e01202 (2013).
27. Reyes, A. *et al.* Gut DNA viromes of Malawian twins discordant for severe acute malnutrition. *Proc. Natl Acad. Sci. USA* **112**, 11941–11946 (2015).
28. Smith, M. I. *et al.* Gut microbiomes of Malawian twin pairs discordant for kwashiorkor. *Science* **339**, 548–554 (2013).
29. Ridaura, V. K. *et al.* Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science* **341**, 1241214 (2013).
30. Jenq, R. R. *et al.* Regulation of intestinal inflammation by microbiota following allogeneic bone marrow transplantation. *J. Exp. Med.* **209**, 903–911 (2012).
31. Davis-Richardson, A. G. *et al.* *Bacteroides dorei* dominates gut microbiome prior to autoimmunity in Finnish children at high risk for type 1 diabetes. *Front. Microbiol.* **5**, 678 (2014).
32. Khosravi, A. *et al.* Gut microbiota promote hematopoiesis to control bacterial infection. *Cell Host Microbe* **15**, 374–381 (2014).
33. Hsiao, E. Y. *et al.* Microbiota modulate behavioral and physiological abnormalities associated with neurodevelopmental disorders. *Cell* **155**, 1451–1463 (2013).  
**With refs 34–36, demonstrates the need to consider the effects of the metagenome and microbiome of all organs including the brain.**
34. Yano, J. M. *et al.* Indigenous bacteria from the gut microbiota regulate host serotonin biosynthesis. *Cell* **161**, 264–276 (2015).
35. Erny, D. *et al.* Host microbiota constantly control maturation and function of microglia in the CNS. *Nat. Neurosci.* **18**, 965–977 (2015).
36. Sampson, T. R. & Mazmanian, S. K. Control of brain development, function, and behavior by the microbiome. *Cell Host Microbe* **17**, 565–576 (2015).
37. Vétizou, M. *et al.* Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. *Science* **350**, 1079–1084 (2015).
38. Sjogren, K. *et al.* The gut microbiota regulates bone mass in mice. *J. Bone Miner. Res.* **27**, 1357–1367 (2012).
39. De Vlamincq, I. *et al.* Temporal response of the human virome to immunosuppression and antiviral therapy. *Cell* **155**, 1178–1187 (2013).
40. Arpaia, N. *et al.* Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. *Nature* **504**, 451–455 (2013).  
**With refs 41, 42, 43 and 47, shows the diverse effects and mechanisms of microbial metabolites (here short chain fatty acids) and emphasizes that mechanisms underlying the metagenome and microbiome require investigation in multiple cell types.**
41. Atarashi, K. *et al.* T<sub>reg</sub> induction by a rationally selected mixture of *Clostridia* strains from the human microbiota. *Nature* **500**, 232–236 (2013).
42. Sefik, E. *et al.* Individual intestinal symbionts induce a distinct population of ROR $\gamma$ <sup>+</sup> regulatory T cells. *Science* (2015).
43. Smith, P. M. *et al.* The microbial metabolites, short-chain fatty acids, regulate colonic T<sub>reg</sub> cell homeostasis. *Science* **341**, 569–573 (2013).
44. Kernbauer, E., Ding, Y. & Cadwell, K. An enteric virus can functionally replace the beneficial cues provided by commensal bacteria. *Nature* **516**, 94–98 (2014).
45. Reinhardt, C. *et al.* Tissue factor and PAR1 promote microbiota-induced intestinal vascular remodelling. *Nature* **483**, 627–631 (2012).
46. Stappenbeck, T. S., Hooper, L. V. & Gordon, J. I. Developmental regulation of intestinal angiogenesis by indigenous microbes via Paneth cells. *Proc. Natl Acad. Sci. USA* **99**, 15451–15455 (2002).
47. Kaiko, G. E. *et al.* The colonic crypt protects stem cells from microbiota-derived metabolites. *Cell* **165**, 1–13 (2016).
48. Osborne, L. C. *et al.* Virus-helminth coinfection reveals a microbiota-independent mechanism of immunomodulation. *Science* **345**, 578–582 (2014).
49. Reese, T. A. *et al.* Helminth infection reactivates latent  $\gamma$ -herpesvirus via cytokine competition at a viral promoter. *Science* **345**, 573–577 (2014).  
**Emphasizes that phenotypes driven by the metagenome can be polymicrobial and are not exclusively bacterial; here viral and protozoan interactions are important.**
50. Wang, Z. *et al.* Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **472**, 57–63 (2011).
51. Arrieta, M. C. *et al.* Early infancy microbial and metabolic alterations affect risk of childhood asthma. *Sci. Transl. Med.* **7**, 307ra152 (2015).
52. Thorburn, A. N. *et al.* Evidence that asthma is a developmental origin disease influenced by maternal diet and bacterial metabolites. *Nat. Commun.* **6**, 7320 (2015).  
**With ref. 51, shows that intestinal metabolites can affect disease pathogenesis in distant organs (here asthma models in the lung).**
53. Buffie, C. G. *et al.* Precision microbiome reconstitution restores bile acid mediated resistance to *Clostridium difficile*. *Nature* **517**, 205–208 (2015).
54. Reese, T. A. *et al.* Sequential infection with common pathogens promotes human-like immune gene expression and altered vaccine response. *Cell Host Microbe* **19**, 1–7 (2016).  
**This report emphasizes the history of microbial interactions with an organism additionally shapes phenotypes and this needs to be modelled in mouse experiments.**
55. McCafferty, J. *et al.* Stochastic changes over time and not founder effects drive cage effects in microbial community assembly in a mouse model. *ISME J.* **7**, 2116–2125 (2013).
56. Pantoja-Feliciano, I. G. *et al.* Biphasic assembly of the murine intestinal microbiota during early development. *ISME J.* **7**, 1112–1115 (2013).
57. Farkas, A. M. *et al.* Induction of Th17 cells by segmented filamentous bacteria in the murine intestine. *J. Immunol. Methods* **421**, 104–111 (2015).
58. Campbell, J. H. *et al.* Host genetic and environmental effects on mouse intestinal microbiota. *ISME J.* **6**, 2033–2044 (2012).
59. Yurkovetskiy, L. *et al.* Gender bias in autoimmunity is influenced by microbiota. *Immunity* **39**, 400–412 (2013).
60. Markle, J. G. *et al.* Sex differences in the gut microbiome drive hormone-dependent regulation of autoimmunity. *Science* **339**, 1084–1088 (2013).
61. Langille, M. G. *et al.* Microbial shifts in the aging mouse gut. *Microbiome* **2**, 50 (2014).
62. Mukherji, A., Kobiita, A., Ye, T. & Chabon, P. Homeostasis in intestinal epithelium is orchestrated by the circadian clock and microbiota cues transduced by TLRs. *Cell* **153**, 812–827 (2013).
63. Thaiss, C. A. *et al.* Transkingdom control of microbiota diurnal oscillations promotes metabolic homeostasis. *Cell* **159**, 514–529 (2014).
64. McInnes, E. F. *et al.* Prevalence of viral, bacterial and parasitological diseases in rats and mice used in research environments in Australasia over a 5-y period. *Lab Anim. (NY)* **40**, 341–350 (2011).
65. Karp, C. L. Unstressing intertemperate models: how cold stress undermines mouse modeling. *J. Exp. Med.* **209**, 1069–1074 (2012).
66. Gaskill, B. N. *et al.* Heat or insulation: behavioral titration of mouse preference for warmth or access to a nest. *PLoS One* **7**, e32799 (2012).
67. Huang, E. Y. *et al.* Using corticosteroids to reshape the gut microbiome: implications for inflammatory bowel diseases. *Inflamm. Bowel Dis.* **21**, 963–972 (2015).
68. Chevalier, C. *et al.* Gut microbiota orchestrates energy homeostasis during cold. *Cell* **163**, 1360–1374 (2015).
69. Wolf, K. J. *et al.* Consumption of acidic water alters the gut microbiome and decreases the risk of diabetes in NOD mice. *J. Histochem. Cytochem.* **62**, 237–250 (2014).



70. Hall, J. E., White, W. J. & Lang, C. M. Acidification of drinking water: its effects on selected biologic phenomena in male mice. *Lab. Anim. Sci.* **30**, 643–651 (1980).
71. Carmody, R. N. *et al.* Diet dominates host genotype in shaping the murine gut microbiota. *Cell Host Microbe* **17**, 72–84 (2015).  
**One example of many cited here that shows in detail the potential rapid effects on environmental influences such as diet on the composition of the microbiome and emphasizes the need to report environmental variables in all experiments.**
72. Jansson, N. *et al.* Gut microbiota facilitates dietary heme-induced epithelial hyperproliferation by opening the mucus barrier in colon. *Proc. Natl Acad. Sci. USA* **112**, 10038–10043 (2015).
73. Zenewicz, L. A. *et al.* IL-22 deficiency alters colonic microbiota to be transmissible and colitogenic. *J. Immunol.* **190**, 5306–5312 (2013).
74. Elinav, E. *et al.* NLRP6 inflammasome regulates colonic microbial ecology and risk for colitis. *Cell* **145**, 745–757 (2011).
75. Garrett, W. S. *et al.* Communicable ulcerative colitis induced by T-bet deficiency in the innate immune system. *Cell* **131**, 33–45 (2007).
76. Olszak, T. *et al.* Microbial exposure during early life has persistent effects on natural killer T cell function. *Science* **336**, 489–493 (2012).  
**With ref. 76, emphasizes the early life effects of the microbiota that must be considered in experimental design to control for the metagenome and microbiome.**
78. Trompette, A. *et al.* Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis. *Nat. Med.* **20**, 159–166 (2014).
79. Sun, J. *et al.* Pancreatic  $\beta$ -cells limit autoimmune diabetes via an immunoregulatory antimicrobial peptide expressed under the influence of the gut Microbiota. *Immunity* **43**, 304–317 (2015).
80. Abt, M. C. *et al.* Commensal bacteria calibrate the activation threshold of innate antiviral immunity. *Immunity* **37**, 158–170 (2012).
81. Ichinohe, T. *et al.* Microbiota regulates immune defense against respiratory tract influenza A virus infection. *Proc. Natl Acad. Sci. USA* **108**, 5354–5359 (2011).
82. Ryu, S. H. & Stappenbeck, T. S. Gut–pancreatic axis AMPLified in islets of Langerhans. *Immunity* **43**, 216–218 (2015).
83. Levy, M. *et al.* Microbiota-modulated metabolites shape the intestinal microenvironment by regulating NLRP6 inflammasome signaling. *Cell* **163**, 1428–1443 (2015).
84. Eberl, G. Addressing the experimental variability associated with the microbiota. *Mucosal Immunol.* **8**, 487–490 (2015).
85. Macpherson, A. J. & McCoy, K. D. Standardised animal models of host microbial mutualism. *Mucosal Immunol.* **8**, 476–486 (2015).  
**Excellent review on the methodology and use of isobiotic mice in experiments.**
86. Goodman, A. L. *et al.* Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. *Proc. Natl Acad. Sci. USA* **108**, 6252–6257 (2011).
87. Atkinson, M. A. & Chervonsky, A. Does the gut microbiota have a role in type 1 diabetes? Early evidence from humans and animal models of the disease. *Diabetologia* **55**, 2868–2877 (2012).
88. Pozzilli, P., Signore, A., Williams, A. J. & Beales, P. E. NOD mouse colonies around the world—recent facts and figures. *Immunol. Today* **14**, 193–196 (1993).
89. Turnbaugh, P. J. *et al.* The human microbiome project. *Nature* **449**, 804–810 (2007).
90. Wheeler, M. L. & Underhill, D. M. Time to cast a larger net. *Nat. Immunol.* **15**, 1000–1001 (2014).
91. Kostic, A. D. *et al.* The dynamics of the human infant gut microbiome in development and in progression toward type 1 diabetes. *Cell Host Microbe* **17**, 260–273 (2015).
92. Hajishengallis, G. *et al.* Low-abundance biofilm species orchestrates inflammatory periodontal disease through the commensal microbiota and complement. *Cell Host Microbe* **10**, 497–506 (2011).  
**With ref. 93, shows that microbiome sequence analysis alone cannot always account for mouse phenotypes; these are examples where low abundance microbes and microbial particles drive phenotypes that would both be missed in sequence analysis.**
93. Hickey, C. A. *et al.* Colitogenic *Bacteroides thetaiotaomicron* antigens access host immune cells in a sulfate-dependent manner via outer membrane vesicles. *Cell Host Microbe* **17**, 672–680 (2015).
94. Segata, N. *et al.* Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat. Methods* **9**, 811–814 (2012).
95. Whitman, W. B., Coleman, D. C. & Wiebe, W. J. Prokaryotes: the unseen majority. *Proc. Natl Acad. Sci. USA* **95**, 6578–6583 (1998).
96. Virgin, H. W., Wherry, E. J. & Ahmed, R. Redefining chronic viral infection. *Cell* **138**, 30–50 (2009).
97. Young, G. R. *et al.* Resurrection of endogenous retroviruses in antibody-deficient mice. *Nature* **491**, 774–778 (2012).
98. Ives, A. *et al.* *Leishmania* RNA virus controls the severity of mucocutaneous leishmaniasis. *Science* **331**, 775–778 (2011).
99. Underhill, D. M. & Iliev, I. D. The mycobiota: interactions between commensal fungi and the host immune system. *Nat. Rev. Immunol.* **14**, 405–416 (2014).
100. Fletcher, S. M., Stark, D., Harkness, J. & Ellis, J. Enteric protozoa in the developed world: a public health perspective. *Clin. Microbiol. Rev.* **25**, 420–449 (2012).
101. Stelekati, E. & Wherry, E. J. Chronic bystander infections and immunity to unrelated antigens. *Cell Host Microbe* **12**, 458–469 (2012).
102. Pfeiffer, J. K. & Virgin, H. W. Viral immunity. Transkingdom control of viral infection and immunity in the mammalian intestine. *Science* **351**, 239–245 (2016).
103. Dantas, G. & Sommer, M. O. Context matters – the complex interplay between resistome genotypes and resistance phenotypes. *Curr. Opin. Microbiol.* **15**, 577–582 (2012).
104. Stern, A. & Sorek, R. The phage–host arms race: shaping the evolution of microbes. *BioEssays* **33**, 43–51 (2011).
105. Brüssow, H., Chanchaya, C. & Hardt, W.-D. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol. Mol. Biol. Rev.* **68**, 560–602 (2004).
106. Fortier, L.-C. & Sekulovic, O. Importance of prophages to evolution and virulence of bacterial pathogens. *Virulence* **4**, 354–365 (2013).
107. Zhang, B. *et al.* Viral infection. Prevention and cure of rotavirus infection via TLR5/NLRC4-mediated production of IL-22 and IL-18. *Science* **346**, 861–865 (2014).
108. Baldridge, M. T. *et al.* Commensal microbes and interferon- $\lambda$  determine persistence of enteric murine norovirus infection. *Science* **347**, 266–269 (2015).
109. Robinson, C. M., Jesudhasan, P. R. & Pfeiffer, J. K. Bacterial lipopolysaccharide binding enhances virion stability and promotes environmental fitness of an enteric virus. *Cell Host Microbe* **15**, 36–46 (2014).
110. Kuss, S. K. *et al.* Intestinal microbiota promote enteric virus replication and systemic pathogenesis. *Science* **334**, 249–252 (2011).
111. Kane, M. *et al.* Successful transmission of a retrovirus depends on the commensal microbiota. *Science* **334**, 245–249 (2011).
112. Barton, E. S. *et al.* Herpesvirus latency confers symbiotic protection from bacterial infection. *Nature* **447**, 326–329 (2007).
113. Furman, D. *et al.* Cytomegalovirus infection enhances the immune response to influenza. *Sci. Transl. Med.* **7**, 281ra43 (2015).
114. Oh, J. Z. *et al.* TLR5-mediated sensing of gut microbiota is necessary for antibody responses to seasonal influenza vaccination. *Immunity* **41**, 478–492 (2014).
115. Williams, W. B. *et al.* HIV-1 vaccines. Diversion of HIV-1 vaccine-induced immunity by gp41-microbiota cross-reactive antibodies. *Science* **349**, aab1253 (2015).
116. Nam, Y.-D. *et al.* Bacterial, archaeal, and eukaryal diversity in the intestines of Korean people. *J. Microbiol.* **46**, 491–501 (2008).
117. Scanlan, P. D. & Marchesi, J. R. Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces. *ISME J.* **2**, 1183–1193 (2008).
118. Elliott, D. E. & Weinstock, J. V. Helminth–host immunological interactions: prevention and control of immune-mediated diseases. *Ann. NY Acad. Sci.* **1247**, 83–96 (2012).
119. Hoffmann, C. *et al.* Archaea and fungi of the human gut microbiome: correlations with diet and bacterial residents. *PLoS One* **8**, e66019 (2013).
120. Kilkenny, C., Browne, W. J., Cuthill, I. C., Emerson, M. & Altman, D. G. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *J. Pharmacol. Pharmacother.* **1**, 94–99 (2010).

**Author Contributions** Both authors contributed equally to planning and writing the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.S.S. ([stappenb@pathology.wustl.edu](mailto:stappenb@pathology.wustl.edu)) and H.W.V. ([virgin@wustl.edu](mailto:virgin@wustl.edu)).

# The genetic history of Ice Age Europe

Qiaomei Fu<sup>1,2,3</sup>, Cosimo Posth<sup>4,5\*</sup>, Mateja Hajdinjak<sup>3\*</sup>, Martin Petr<sup>3</sup>, Swapan Mallick<sup>2,6,7</sup>, Daniel Fernandes<sup>8,9</sup>, Anja Furtwängler<sup>4</sup>, Wolfgang Haak<sup>5,10</sup>, Matthias Meyer<sup>3</sup>, Alissa Mittnik<sup>4,5</sup>, Birgit Nickel<sup>3</sup>, Alexander Peltzer<sup>4</sup>, Nadin Rohland<sup>2</sup>, Viviane Slon<sup>3</sup>, Sahra Talamo<sup>11</sup>, Iosif Lazaridis<sup>2</sup>, Mark Lipson<sup>2</sup>, Iain Mathieson<sup>2</sup>, Stephan Schiffels<sup>5</sup>, Pontus Skoglund<sup>2</sup>, Anatoly P. Derevianko<sup>12,13</sup>, Nikolai Drozdov<sup>12</sup>, Vyacheslav Slavinsky<sup>12</sup>, Alexander Tsybankov<sup>12</sup>, Renata Grifoni Cremonesi<sup>14</sup>, Francesco Mallegni<sup>15</sup>, Bernard Gély<sup>16</sup>, Eligio Vacca<sup>17</sup>, Manuel R. González Morales<sup>18</sup>, Lawrence G. Straus<sup>18,19</sup>, Christine Neugebauer-Maresch<sup>20</sup>, Maria Teschler-Nicola<sup>21,22</sup>, Silviu Constantin<sup>23</sup>, Oana Teodora Moldovan<sup>24</sup>, Stefano Benazzi<sup>11,25</sup>, Marco Peresani<sup>26</sup>, Donato Coppola<sup>27,28</sup>, Martina Lari<sup>29</sup>, Stefano Ricci<sup>30</sup>, Annamaria Ronchitelli<sup>30</sup>, Frédérique Valentin<sup>31</sup>, Corinne Thevenet<sup>32</sup>, Kurt Wehrberger<sup>33</sup>, Dan Grigorescu<sup>34</sup>, Hélène Rougier<sup>35</sup>, Isabelle Crevecoeur<sup>36</sup>, Damien Flas<sup>37</sup>, Patrick Semal<sup>38</sup>, Marcello A. Mannino<sup>11,39</sup>, Christophe Cupillard<sup>40,41</sup>, Hervé Bocherens<sup>42,43</sup>, Nicholas J. Conard<sup>43,44</sup>, Katerina Harvati<sup>43,45</sup>, Vyacheslav Moiseyev<sup>46</sup>, Dorothee G. Drucker<sup>42</sup>, Jiří Svoboda<sup>47,48</sup>, Michael P. Richards<sup>11,49</sup>, David Caramelli<sup>29</sup>, Ron Pinhasi<sup>8</sup>, Janet Kelso<sup>3</sup>, Nick Patterson<sup>6</sup>, Johannes Krause<sup>4,5,43§</sup>, Svante Pääbo<sup>3§</sup> & David Reich<sup>2,6,7§</sup>

**Modern humans arrived in Europe ~45,000 years ago, but little is known about their genetic composition before the start of farming ~8,500 years ago. Here we analyse genome-wide data from 51 Eurasians from ~45,000–7,000 years ago. Over this time, the proportion of Neanderthal DNA decreased from 3–6% to around 2%, consistent with natural selection against Neanderthal variants in modern humans. Whereas there is no evidence of the earliest modern humans in Europe contributing to the genetic composition of present-day Europeans, all individuals between ~37,000 and ~14,000 years ago descended from a single founder population which forms part of the ancestry of present-day Europeans. An ~35,000-year-old individual from northwest Europe represents an early branch of this founder population which was then displaced across a broad region, before reappearing in southwest Europe at the height of the last Ice Age ~19,000 years ago. During the major warming period after ~14,000 years ago, a genetic component related to present-day Near Easterners became widespread in Europe. These results document how population turnover and migration have been recurring themes of European prehistory.**

Modern humans arrived in Europe around 45,000 years ago and have lived there ever since, even during the Last Glacial Maximum 25,000–19,000 years ago when large parts of Europe were covered in ice<sup>1</sup>. A major question is how climatic fluctuations influenced the population history of Europe and to what extent changes in material cultures documented by archaeology corresponded to movements of people. To date, it has been difficult to address this question because genome-wide ancient DNA has been retrieved from just four Upper Palaeolithic

individuals from Europe<sup>2–4</sup>. Here we assemble and analyse genome-wide data from 51 modern humans dating from 45,000 to 7,000 years ago (Extended Data Table 1; Supplementary Information section 1).

## Ancient DNA retrieval

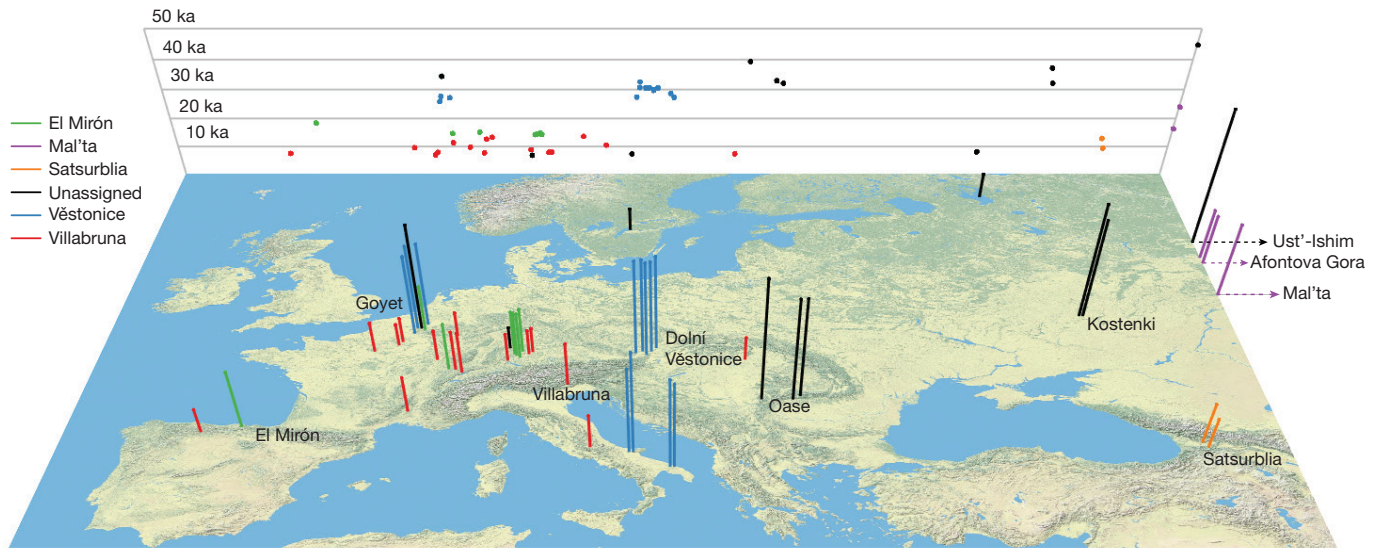
We extracted DNA from human remains in dedicated clean rooms<sup>5</sup>, and transformed the extracts into Illumina sequencing libraries<sup>6–8</sup>. A major challenge in ancient DNA research is that the vast majority

<sup>1</sup>Key Laboratory of Vertebrate Evolution and Human Origins of Chinese Academy of Sciences, IVPP, CAS, Beijing 100044, China. <sup>2</sup>Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>3</sup>Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany. <sup>4</sup>Institute for Archaeological Sciences, Archaeology and Palaeogenetics, University of Tübingen, 72070 Tübingen, Germany. <sup>5</sup>Department of Archaeogenetics, Max Planck Institute for the Science of Human History, 07745 Jena, Germany. <sup>6</sup>Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA. <sup>7</sup>Howard Hughes Medical Institute, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>8</sup>School of Archaeology and Earth Institute, University College Dublin, Belfield, Dublin 4, Ireland. <sup>9</sup>CIAS, Department of Life Sciences, University of Coimbra, 3000-456 Coimbra, Portugal. <sup>10</sup>Australian Centre for Ancient DNA, School of Biological Sciences, The University of Adelaide, SA-5005 Adelaide, Australia. <sup>11</sup>Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany. <sup>12</sup>Institute of Archaeology and Ethnography, Russian Academy of Sciences, Siberian Branch, 17 Novosibirsk, RU-630090, Russia. <sup>13</sup>Altai State University, Barnaul, RU-656049, Russia. <sup>14</sup>Dipartimento di Civiltà e Forme del Sapere, Università di Pisa, 56126 Pisa, Italy. <sup>15</sup>Department of Biology, University of Pisa, 56126 Pisa, Italy. <sup>16</sup>Direction régionale des affaires culturelles Rhône-Alpes, 69283 Lyon, Cedex 01, France. <sup>17</sup>Dipartimento di Biologia, Università degli Studi di Bari 'Aldo Moro', 70125 Bari, Italy. <sup>18</sup>Instituto Internacional de Investigaciones Prehistóricas, Universidad de Cantabria, 39005 Santander, Spain. <sup>19</sup>Department of Anthropology, MSC01 1040, University of New Mexico, Albuquerque, New Mexico 87131-0001, USA. <sup>20</sup>Quaternary Archaeology, Institute for Oriental and European Archaeology, Austrian Academy of Sciences, 1010 Vienna, Austria. <sup>21</sup>Department of Anthropology, Natural History Museum Vienna, 1010 Vienna, Austria. <sup>22</sup>Department of Anthropology, University of Vienna, 1090 Vienna, Austria. <sup>23</sup>'Emil Racoviță' Institute of Speleology, 010986 Bucharest 12, Romania. <sup>24</sup>'Emil Racoviță' Institute of Speleology, Cluj Branch, 400006 Cluj, Romania. <sup>25</sup>Department of Cultural Heritage, University of Bologna, 48121 Ravenna, Italy. <sup>26</sup>Sezione di Scienze Preistoriche e Antropologiche, Dipartimento di Studi Umanistici, Università di Ferrara, 44100 Ferrara, Italy. <sup>27</sup>Università degli Studi di Bari 'Aldo Moro', 70125 Bari, Italy. <sup>28</sup>Museo di 'Civiltà preclassiche della Murgia meridionale', 72017 Ostuni, Italy. <sup>29</sup>Dipartimento di Biologia, Università di Firenze, 50122 Florence, Italy. <sup>30</sup>Dipartimento di Scienze Fisiche, della Terra e dell'Ambiente, U.R. Preistoria e Antropologia, Università degli Studi di Siena, 53100 Siena, Italy. <sup>31</sup>CNRS/UMR 7041 ArScAn MAE, 92023 Nanterre, France. <sup>32</sup>INRAP/UMR 8215 Trajectoires 21, 92023 Nanterre, France. <sup>33</sup>Ulmer Museum, 89073 Ulm, Germany. <sup>34</sup>University of Bucharest, Faculty of Geology and Geophysics, Department of Geology, 01041 Bucharest, Romania. <sup>35</sup>Department of Anthropology, California State University Northridge, Northridge, California 91330-8244, USA. <sup>36</sup>Université de Bordeaux, CNRS, UMR 5199-PACEA, 33615 Pessac Cedex, France. <sup>37</sup>TRACES – UMR 5608, Université Toulouse Jean Jaurès, Maison de la Recherche, 31058 Toulouse Cedex 9, France. <sup>38</sup>Royal Belgian Institute of Natural Sciences, 1000 Brussels, Belgium. <sup>39</sup>Department of Archaeology, School of Culture and Society, Aarhus University, 8270 Højbjerg, Denmark. <sup>40</sup>Service Régional d'Archéologie de Franche-Comté, 25043 Besançon Cedex, France. <sup>41</sup>Laboratoire Chronoenvironnement, UMR 6249 du CNRS, UFR des Sciences et Techniques, 25030 Besançon Cedex, France. <sup>42</sup>Department of Geosciences, Biogeology, University of Tübingen, 72074 Tübingen, Germany. <sup>43</sup>Senckenberg Centre for Human Evolution and Palaeoenvironment, University of Tübingen, 72072 Tübingen, Germany. <sup>44</sup>Department of Early Prehistory and Quaternary Ecology, University of Tübingen, 72070 Tübingen, Germany. <sup>45</sup>Institute for Archaeological Sciences, Paleoanthropology, University of Tübingen, 72070 Tübingen, Germany. <sup>46</sup>Museum of Anthropology and Ethnography, Saint Petersburg 34, Russia. <sup>47</sup>Department of Anthropology, Faculty of Science, Masaryk University, 611 37 Brno, Czech Republic. <sup>48</sup>Institute of Archaeology at Brno, Academy of Science of the Czech Republic, 69129 Dolní Věstonice, Czech Republic. <sup>49</sup>Department of Archaeology, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada.

\*These authors contributed equally to this work.

§These authors jointly supervised this work.





**Figure 1 | Location and age of the 51 ancient modern humans.** Each bar corresponds to an individual, the colour code designates the genetically defined cluster of individuals, and the height is proportional to age (the background grid shows a projection of longitude against age). To help in

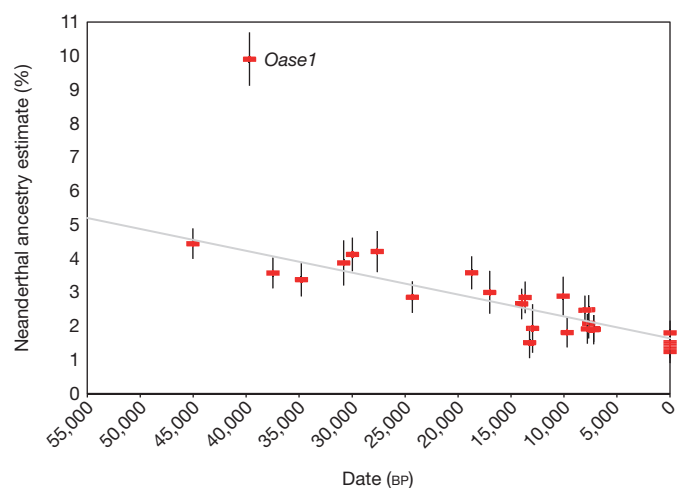
visualization, we add jitter for sites with multiple individuals from nearby locations. Four individuals from Siberia are plotted at the far eastern edge of the map. ka, thousand years ago.

of the DNA extracted from most specimens is of microbial origin, making random shotgun sequencing prohibitively expensive. We addressed this problem by enriching the libraries for between 390,000 and 3.7 million single nucleotide polymorphisms (SNPs) in the nuclear genome via hybridizing to pools of previously synthesized 52-base-pair oligonucleotide probes targeting these positions. This makes it possible to generate genome-wide data from samples with high percentages of microbial DNA that are not practical to study by shotgun sequencing<sup>3,9</sup>. We sequenced the isolated DNA fragments from both ends, and mapped the consensus sequences to the human genome (hg19), retaining fragments that overlapped the targeted SNPs. After removing fragments with identical start and end positions to eliminate duplicates produced during library amplification, we chose one fragment at random to represent each individual at each SNP.

Contamination from present-day human DNA is a danger in ancient DNA research. To address this, we took advantage of three characteristic features of ancient DNA (Supplementary Information section 2). First, for an uncontaminated specimen, we expect only a single mitochondrial DNA sequence to be present, allowing us to detect contamination as a mixture of mitochondrial sequences. Second, because males carry a single X chromosome, we can detect contamination in male specimens as polymorphisms on chromosome X<sup>10</sup>. Third, cytosines at the ends of genuine ancient DNA molecules are often deaminated, resulting in apparent cytosine to thymine substitutions<sup>11</sup>, and thus we can filter out contaminating molecules by restricting analysis to those with evidence of such deamination<sup>12</sup>. For libraries from males with evidence of mitochondrial DNA contamination or X chromosomal contamination estimates >2.5%—as well as for all libraries from females—we restricted the analyses to sequences with evidence of cytosine deamination (Supplementary Information section 2). After merging libraries from the same individual and limiting to individuals with >4,000 targeted SNPs covered at least once, 38 individuals remained, which we merged with newly generated shotgun sequencing data from the *Karelia* individual<sup>9</sup> (2.0-fold coverage), and published data from ancient<sup>2–4,7,13–19</sup> and present-day humans<sup>20</sup>. The final data set includes 51 ancient modern humans, of which 16 had at least 790,000 SNPs covered (Fig. 1; Extended Data Table 1).

### Natural selection reduced Neanderthal ancestry over time

We used two previously published statistics<sup>3,7,21</sup> to test if the proportion of Neanderthal ancestry in Eurasians changed over the last 45,000 years. Whereas on the order of 2% of present-day Eurasian DNA is of Neanderthal origin (Extended Data Table 2), the ancient modern human genomes carry significantly more Neanderthal DNA (Fig. 2) ( $P < 10^{-12}$ ). Using one statistic, we estimate a decline from 4.3–5.7% from a time shortly after introgression to 1.1–2.2% in Eurasians today (Fig. 2). Using the other statistic, we estimate a decline from 3.2–4.2% to 1.8–2.3% (Extended Data Fig. 1 and Extended Data Table 3). Because all of the European individuals we analysed dating to between 37,000 and 14,000 years ago are consistent with descent from a single founding



**Figure 2 | Decrease of Neanderthal ancestry over time.** Plot of radiocarbon date against Neanderthal ancestry for individuals with at least 200,000 SNPs covered, along with present-day Eurasians (standard errors are from a block jackknife). The least squares fit (grey) excludes the data from *Oase1* (an outlier with recent Neanderthal ancestry) and three present-day European populations (known to have less Neanderthal ancestry than east Asians). The slope is significantly negative for all eleven subsets of individuals we analysed ( $10^{-29} < P < 10^{-11}$  based on a block jackknife) (Extended Data Table 3). BP, before present.

population, admixture with populations with lower Neanderthal ancestry cannot explain the steady decrease in Neanderthal-derived DNA that we detect during this period, showing that natural selection against Neanderthal DNA must have driven this phenomenon (Fig. 2). We also obtained an independent line of evidence for selection from our observation that the decrease in Neanderthal-derived alleles is more marked near genes than in less constrained regions of the genome ( $P = 0.010$ ) (Extended Data Table 3; Supplementary Information section 3)<sup>22–25</sup>.

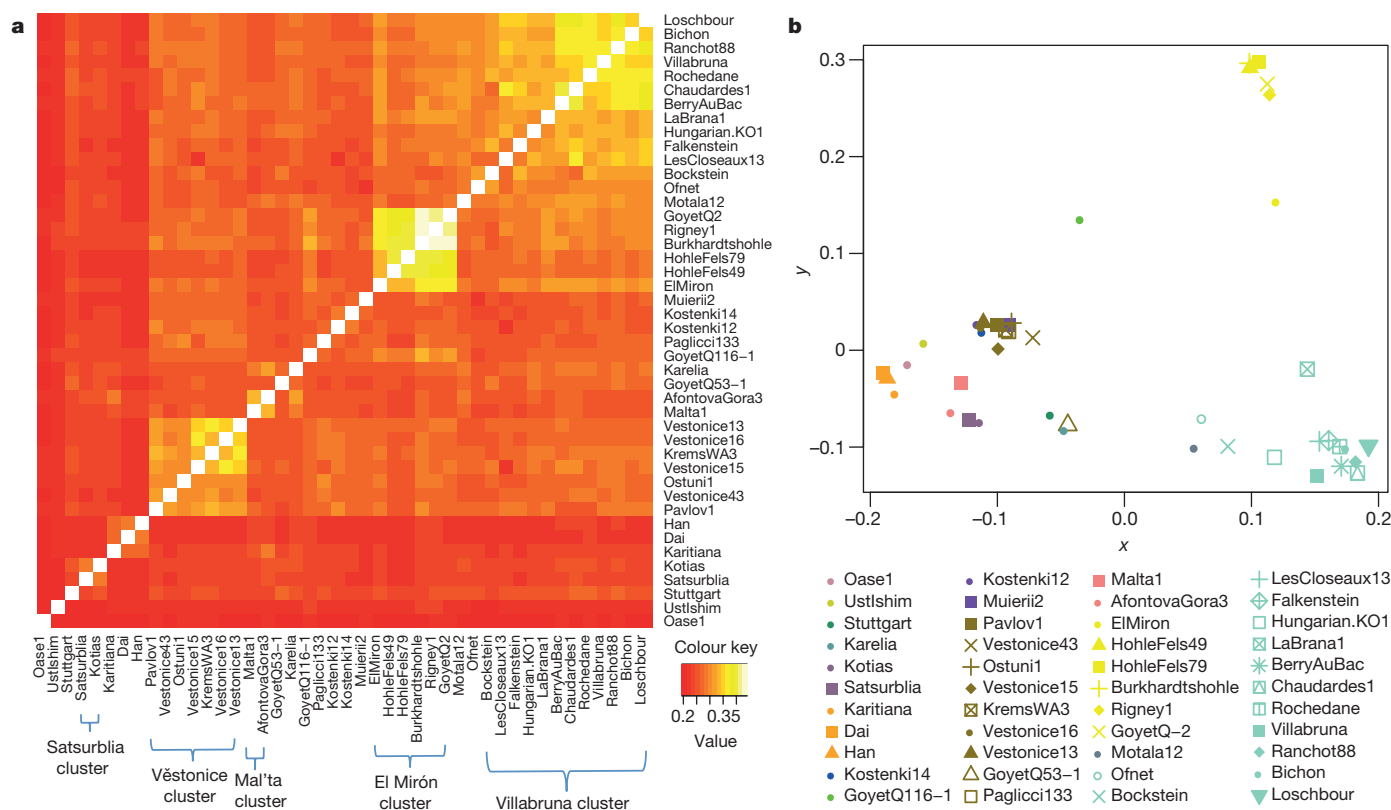
### Chromosome Y, mtDNA, and significant mutations

We used the proportion of sequences mapping to the Y chromosome to infer sex (Extended Data Table 4; Supplementary Information section 4), and determined Y chromosome haplogroups for the males. We were surprised to find haplogroup R1b in the ~14,000-year-old *Villabruna* individual from Italy. While the predominance of R1b in western Europe today owes its origin to Bronze Age migrations from the eastern European steppe<sup>9</sup>, its presence in *Villabruna* and in a ~7,000-year-old farmer from Iberia<sup>9</sup> documents a deeper history of this haplotype in more western parts of Europe. Additional evidence of an early link between West and East comes from the *HERC2* locus, where a derived allele that is the primary driver of light eye colour in Europeans appears nearly simultaneously in specimens from Italy and the Caucasus ~14,000–13,000 years ago. Extended Data Table 5 presents results for additional alleles of biological importance. When analysing the mitochondrial genomes we noted the presence of haplogroup M in a ~27,000-year-old individual from southern Italy (*Ostuni1*) in agreement with the observation that this haplogroup, which today occurs in Asia and is absent in Europe, was

present in pre-Last Glacial Maximum Europe and was subsequently lost<sup>26</sup>. We also find that the ~33,000-year-old *Muierii2* from Romania carries a basal version of haplogroup U6, in agreement with the hypothesis that the presence of derived versions of this haplogroup in North Africans today is due to back-migration from western Eurasia<sup>27</sup>.

## Genetic clustering of the ancient specimens

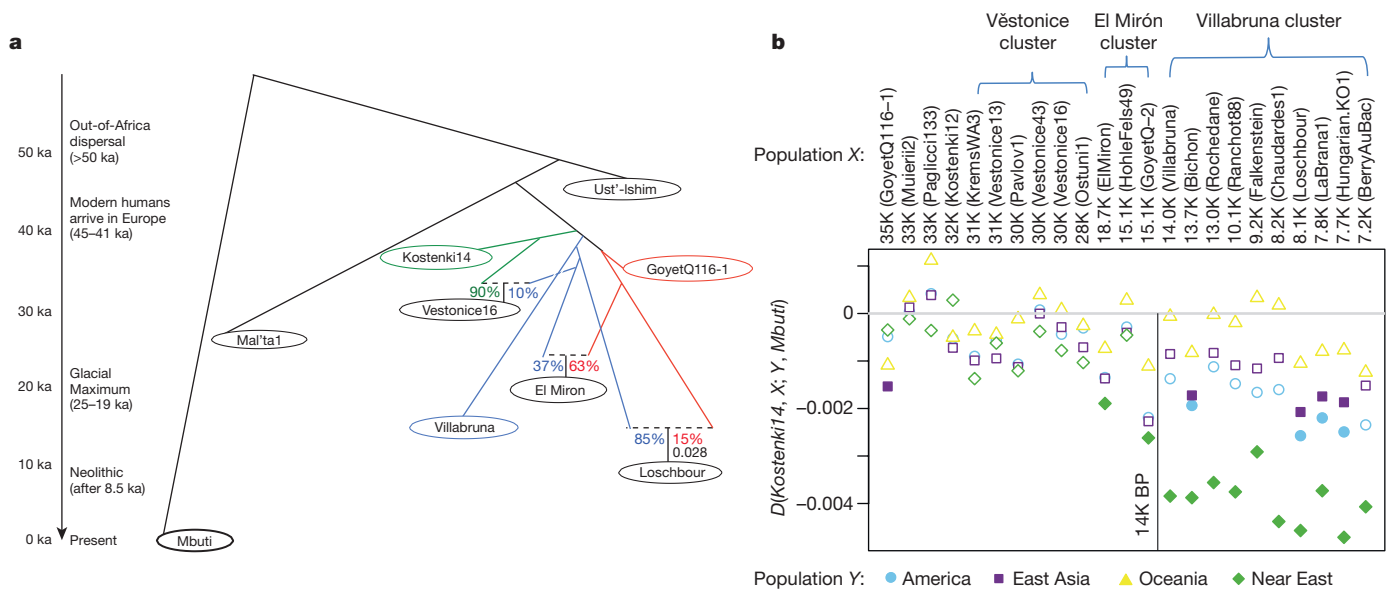
This data set provides an unprecedented opportunity to study the population history of Upper Palaeolithic Europe over more than 30,000 years. In order not to prejudice any association between genetic and archaeological groupings among the individuals studied, we first allowed the genetic data alone to drive the groupings of the specimens, and only afterward examined their associations with archaeological cultural complexes. We began by computing  $f_3$ -statistics<sup>14</sup> of the form  $f_3(X, Y; Mbuti)$ , which measure shared genetic drift between a pair of ancient individuals after divergence from an outgroup (here *Mbuti* from sub-Saharan Africa) (Fig. 3a and Extended Data Fig. 2). Through multi-dimensional scaling (MDS) analysis of this matrix (Fig. 3b), as well as through  $D$ -statistic analyses<sup>28</sup> (Supplementary Information section 5), we identify five clusters of individuals who share substantial amounts of genetic drift. We name these clusters after the oldest individual in each cluster with >1.0-fold coverage (Supplementary Information section 5; Extended Data Table 1). In contrast, we were not able to identify clear structure among the individuals studied based on model-based clustering<sup>29,30</sup>, which may reflect the fact that many of the individuals are so ancient that present-day human variation is not very relevant to understanding their patterns of genetic differentiation<sup>4,13</sup>. The ‘Vestonice Cluster’ is composed of 14 pre-Last



**Figure 3 | Genetic clustering of the ancient modern humans.** **a**, Shared genetic drift measured by  $f_3(X, Y; Mbuti)$  among individuals with at least 30,000 SNPs covered (for *AfontovaGora3*, *ElMiron*, *Falkenstein*, *GoyetQ-2*, *GoyetQ53-1*, *HohleFels49*, *HohleFels79*, *LesCloseaux13*, *Ofnet*, *Ranchot88* and *Rigney1*, we use all sequences for higher resolution). Lighter colours indicate more shared drift. **b**, Multi-dimensional scaling (MDS) analysis,

computed using the R software *cmdscales* package, highlights the main genetic groupings analysed in this study: Věstonice Cluster (brown), Mal'ta Cluster (pink), El Mirón Cluster (yellow), Villabruna Cluster (light green), and Satsurblia Cluster (dark purple). The affinity of *GoyetQ116-1* (dark green) to the El Mirón Cluster is evident in both views of the data.





**Figure 4 | Population history inferences.** **a**, Admixture graph relating selected high coverage individuals. Dashed lines show inferred admixture events; the estimated mixture proportions fitted using the ADMIXTUREGRAPH software are labelled<sup>28</sup> (the estimated genetic drift on each branch is given in a version of this graph shown in Supplementary Information section 6). The individuals are positioned vertically based on their radiocarbon date, but we caution that the population split times are not accurately known. Colour is used to highlight important early branches of the European founder population: the *Kostenki14* lineage is modelled as the predominant contributor to the Věstovice Cluster

Glacial Maximum individuals from 34,000–26,000 years ago, who are all associated with the archaeologically defined Gravettian culture. The ‘Mal’ta Cluster’ is composed of three individuals who lived between 24,000–17,000 years ago from the Lake Baikal region of Siberia. The ‘El Mirón Cluster’ is composed of seven post-Last Glacial Maximum individuals from 19,000–14,000 years ago, who are all associated with the Magdalenian culture. The ‘Villabruna Cluster’ is composed of 15 post-Last Glacial Maximum individuals from 14,000–7,000 years ago, associated with the Azilian, Epipaleolithic and Mesolithic cultures. The ‘Satsurblia Cluster’ is composed of two individuals from 13,000–10,000 years ago from the southern Caucasus<sup>2</sup>. Ten individuals were not assigned to any cluster, either because they represented distinct early lineages (*Ust’-Ishim*, *Oase1*, *Kostenki14*, *GoyetQ116-1*, *Muierii2*, *Cioclovina1* and *Kostenki12*), because they were admixed between clusters (*Karelia* or *Motala12*), or because they were of very different ancestry (*Stuttgart*). To classify the ancestry of additional low coverage individuals, we built an admixture graph that fits the allele frequency correlation patterns among high-coverage individuals<sup>28</sup> (Fig. 4a; Supplementary Information section 6). We fit each low-coverage individual into the graph in turn, using all DNA fragments from these individuals, rather than just fragments with evidence of cytosine deamination, and account for contamination by modelling (Supplementary Information section 7).

### A founding population for Europeans 37–14 ka

A previous genetic analysis of early modern humans in Europe using data from the ~37,000-year-old *Kostenki14* suggested that the population to which *Kostenki14* belonged harboured within it the three major lineages that exist in mixed form in Europe today<sup>4,15</sup>: (1) a lineage related to all later pre-Neolithic Europeans, (2) a ‘Basal Eurasian’ lineage that split from the ancestors of Europeans and east Asians before they separated from each other; and (3) a lineage related to the ~24,000-year-old *Mal’ta1* from Siberia. With our more extensive sampling of Ice Age Europe, we find no support for this. When we test whether the ~45,000-year-old *Ust’-Ishim*—an early Eurasian without

(green); the *GoyetQ116-1* lineage as the predominant contributor to the El Mirón Cluster (red); and the *Villabruna* lineage as broadly represented across many clusters. **b**, Drawing together of European and Near Eastern populations ~14,000 years ago. Plot of affinity of each pre-Neolithic European population *X* to non-Africans outside Europe *Y* moving forward in time, comparing to *Kostenki14* as a baseline; values  $Z < -3$  standard errors below zero are indicated with filled symbols (we restricted to individuals with >50,000 SNPs). We observe an affinity to Near Easterners beginning with the Villabruna Cluster, and another to east Asians that affects a subset of the Villabruna Cluster.

any evidence of Basal Eurasian ancestry—shares more alleles with one test individual or another by computing statistics of the form  $D(Test_1, Test_2; Ust’-Ishim, Mbuti)$ , we find that the statistic is consistent with zero when the *Test* populations are any pre-Neolithic Europeans or present-day east Asians<sup>3,13</sup>. This would not be expected if some of the pre-Neolithic Europeans, including *Kostenki14*, had Basal Eurasian ancestry (Supplementary Information section 8). We also find no evidence for the suggestion that the *Mal’ta1* lineage contributed to Upper Palaeolithic Europeans<sup>4</sup>, because when we compute the statistic  $D(Test_1, Test_2; Mal’ta1, Mbuti)$ , we find that the statistic is indistinguishable from zero when the *Test* populations are any pre-Neolithic Europeans beginning with *Kostenki14*, consistent with descent from a single founder population since separation from the lineage leading to *Mal’ta1* (Supplementary Information section 9). A corollary of this finding is that the widespread presence of *Mal’ta1*-related ancestry in present-day Europeans<sup>15</sup> is probably explained by migrations from the Eurasian steppe in the Neolithic and Bronze Age periods<sup>9</sup>.

### Resurfacing of a European lineage in the Glacial Maximum

Among the newly reported individuals, *GoyetQ116-1* from present-day Belgium is the oldest at ~35,000 years ago. This individual is similar to the ~37,000-year-old *Kostenki14* and all later individuals in that it shares more alleles with present-day Europeans (for example, *French*) than with east Asians (for example, *Han*). In contrast, *Ust’-Ishim* and *Oase1*, which predate *GoyetQ116-1* and *Kostenki14*, do not show any distinctive affinity to later Europeans (Extended Data Table 6). Thus, from about 37,000 years ago, populations in Europe shared at least some ancestry with present Europeans. However, *GoyetQ116-1* differs from *Kostenki14* and from all individuals of the succeeding Věstovice Cluster in that both  $f_3$ -statistics (Fig. 3; Extended Data Fig. 2) and  $D$ -statistics show that it shares more alleles with members of the El Mirón Cluster who lived 19,000–14,000 years ago than with other pre-Neolithic Europeans (Supplementary Information section 10). Thus, *GoyetQ116-1* has an affinity to individuals who lived more than 15,000 years later. While at least half of the ancestry of all

El Mirón Cluster individuals comes from the lineage represented by *GoyetQ116-1*, this proportion varies among individuals with the largest amount found outside Iberia ( $Z = -4.8$ ) (Supplementary Information section 10).

### Europe and the Near East drew together around 14 ka

Beginning around 14,000 years ago with the Villabruna Cluster, the strong affinity to *GoyetQ116-1* seen in El Mirón Cluster individuals who belong to the Late Glacial Magdalenian culture becomes greatly attenuated (Supplementary Information section 10). To test if this change might reflect gene flow from populations that did not descend from the >37,000-year-old European founder population, we computed statistics of the form  $D(\text{Early European}, \text{Later European}; Y, \text{Mbuti})$  where  $Y$  are various present-day non-Africans. If no gene flow from exogenous populations occurred, this statistic is expected to be zero. Figure 4b shows that it is consistent with zero ( $|Z| < 3$ ) for nearly all individuals dating to between about 37,000 and 14,000 years ago. However, beginning with the Villabruna Cluster, it becomes highly significantly negative in comparisons where the non-European population ( $Y$ ) is Near Easterners (Fig. 4b; Extended Data Fig. 3; Supplementary Information section 11). This must reflect a contribution to the Villabruna Cluster from a lineage also found in present-day Near Easterners (Fig. 4b).

The Satsurblia Cluster individuals from the Caucasus dating to ~13,000–10,000 years ago<sup>2</sup> share more alleles with the Villabruna Cluster individuals than they do with earlier Europeans, indicating that they are related to the population that contributed new alleles to people in the Villabruna Cluster, although they cannot be the direct source of the gene flow. One reason for this is that the Satsurblia Cluster carries large amounts of Basal Eurasian ancestry while Villabruna Cluster individuals do not<sup>2</sup> (Supplementary Information section 12; Extended Data Fig. 4). One possible explanation for the sudden drawing together of the ancestry of Europe and the Near East at this time is long-distance migrations from the Near East into Europe. However, a plausible alternative is population structure, whereby Upper Palaeolithic Europe harboured multiple groups that differed in their relationship to the Near East, with the balance shifting among groups as a result of demographic changes after the Glacial Maximum.

The Villabruna Cluster is represented by the largest number of individuals in this study. This allows us to study heterogeneity within this cluster (Supplementary Information section 13). First, we detect differences in the degree of allele sharing with members of the El Mirón Cluster, as revealed by significant statistics of the form  $D(\text{Test}_1, \text{Test}_2; \text{El Mirón Cluster}, \text{Mbuti})$ . Second, we detect an excess of allele sharing with east Asians in a subset of Villabruna Cluster individuals—beginning with an ~13,000-year-old individual from Switzerland—as revealed by significant statistics of the form  $D(\text{Test}_1, \text{Test}_2; \text{Han}, \text{Mbuti})$  (Fig. 4b and Extended Data Fig. 3). For example, *Han* Chinese share more alleles with two Villabruna Cluster individuals (*Loschbour* and *LaBraná*) than they do with *Kostenki14*, as reflected in significantly negative statistics of the form  $D(\text{Kostenki14}, \text{Loschbour/LaBraná}; \text{Han}, \text{Mbuti})$ <sup>4</sup>. This statistic was originally interpreted as evidence of Basal Eurasian ancestry in *Kostenki14*. However, because this statistic is consistent with zero when *Han* is replaced with *Ust'-Ishim*, these findings cannot be driven by Basal Eurasian ancestry (as we discuss earlier), and must instead be driven by gene flow between populations related to east Asians and the ancestors of some Europeans (Supplementary Information section 8).

### Conclusions

We show that the population history of pre-Neolithic Europe was complex in several respects. First, at least some of the initial modern humans to appear in Eurasia, exemplified by *Ust'-Ishim* and *Oase1*, failed to contribute appreciably to the current European gene pool<sup>3,13</sup>. Only from around 37,000 years ago do all the European individuals analysed share ancestry with present-day Europeans. Second, from

the time of *Kostenki14* about 37,000 years ago until the time of the Villabruna Cluster about 14,000 years ago, all individuals seem to derive from a single ancestral population with no evidence of substantial genetic influx from elsewhere. It is interesting that during this time, the Mal'ta Cluster is not represented in any of the individuals we sampled from Europe. Thus, while individuals assigned to the Gravettian cultural complex in Europe are associated with the Věstonice Cluster, there is no genetic connection between them and the Mal'ta1 individual in Siberia, despite the fact that Venus figurines are associated with both. This suggests that if this similarity is not a coincidence<sup>31</sup>, it reflects diffusion of ideas rather than movements of people. Third, we find that *GoyetQ116-1* derives from a different deep branch of the European founder population than the Věstonice Cluster which became predominant in many places in Europe between 34,000 and 26,000 years ago including at Goyet. *GoyetQ116-1* is chronologically associated with the Aurignacian cultural complex. Thus, the subsequent spread of the Věstonice Cluster shows that the diffusion of the Gravettian cultural complex was mediated at least in part by population movements. Fourth, the population represented by *GoyetQ116-1* did not disappear, as its descendants became widespread again after ~19,000 years ago in the El Mirón Cluster when we detect them in Iberia. The El Mirón Cluster is associated with the Magdalenian culture and may represent a post-Glacial Maximum expansion from southwestern European refugia<sup>32</sup>. Fifth, beginning with the Villabruna Cluster at least ~14,000 years ago, all European individuals analysed show an affinity to the Near East. This correlates in time to the Bølling-Allerød interstadial, the first significant warming period after the Glacial Maximum<sup>33</sup>. Archaeologically, it correlates with cultural transitions within the Epigravettian in southern Europe<sup>34</sup> and the Magdalenian-to-Azilian transition in western Europe<sup>35</sup>. Thus, the appearance of the Villabruna Cluster may reflect migrations or population shifts within Europe at the end of the Ice Age, an observation that is also consistent with the evidence of mitochondrial DNA turnover<sup>26,36</sup>. One scenario that could explain these patterns is a population expansion from southeastern European or west Asian refugia after the Glacial Maximum, drawing together the genetic ancestry of Europe and the Near East. Sixth, within the Villabruna Cluster, some, but not all, individuals have an affinity to east Asians. An important direction for future work will be to generate similar ancient DNA data from southeastern Europe and the Near East to arrive at a more complete picture of the Upper Palaeolithic population history of western Eurasia.

**Online Content** Methods, along with any additional Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 18 December 2015; accepted 12 April 2016.**

**Published online 2 May 2016.**

- Gamble, C., Davies, W., Pettitt, P. & Richards, M. Climate change and evolving human diversity in Europe during the last glacial. *Philosoph. Trans. Royal Soc. B* **359**, 243–253 (2004).
- Jones, E. R. et al. Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nature Commun.* **6**, 8912 (2015).
- Fu, Q. et al. An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524**, 216–219 (2015).
- Seguin-Orlando, A. et al. Paleogenomics. Genomic structure in Europeans dating back at least 36,200 years. *Science* **346**, 1113–1118 (2014).
- Dabney, J. et al. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl Acad. Sci. USA* **110**, 15758–15763 (2013).
- Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols* **2010**, pdb.prot5448 (2010).
- Meyer, M. et al. A high-coverage genome sequence from an archaic Denisovan individual. *Science* **338**, 222–226 (2012).
- Rohland, N., Harney, E., Mallick, S., Nordenfelt, S. & Reich, D. Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. *Phil. Trans. R. Soc. Lond. B* **370**, 20130624 (2015).
- Haak, W. et al. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015).
- Korneliusen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics* **15**, 356 (2014).



11. Krause, J. *et al.* A complete mtDNA genome of an early modern human from Kostenki, Russia. *Curr. Biol.* **20**, 231–236 (2010).
12. Skoglund, P. *et al.* Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science* **336**, 466–469 (2012).
13. Fu, Q. *et al.* Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* **514**, 445–449 (2014).
14. Raghavan, M. *et al.* Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* **505**, 87–91 (2014).
15. Lazaridis, I. *et al.* Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409–413 (2014).
16. Olalde, I. *et al.* Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature* **507**, 225–228 (2014).
17. Gamba, C. *et al.* Genome flux and stasis in a five millennium transect of European prehistory. *Nature Commun.* **5**, 5257 (2014).
18. Green, R. E. *et al.* A draft sequence of the Neandertal genome. *Science* **328**, 710–722 (2010).
19. Prüfer, K. *et al.* The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* **505**, 43–49 (2014).
20. Mallick, S. *et al.* The landscape of human genome diversity. *Nature* (in the press).
21. Reich, D. *et al.* Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* **468**, 1053–1060 (2010).
22. Sankararaman, S. *et al.* The genomic landscape of Neanderthal ancestry in present-day humans. *Nature* **507**, 354–357 (2014).
23. Vernot, B. & Akey, J. M. Resurrecting surviving Neandertal lineages from modern human genomes. *Science* **343**, 1017–1021 (2014).
24. Harris, K. & Nielsen, R. The genetic cost of Neandertal introgression. *Genetics* <http://dx.doi.org/10.1534/genetics.116.186890> (2016).
25. Juric, I., Aeschbacher, S. & Coop, G. The strength of selection against Neandertal introgression. Preprint at <http://dx.doi.org/10.1101/030148> (2015).
26. Posth, C. *et al.* Pleistocene mitochondrial genomes suggest a single major dispersal of Non-Africans and a Late Glacial population turnover in Europe. *Curr. Biol.* **26**, 827–833 (2016).
27. Olivieri, A. *et al.* The mtDNA legacy of the Levantine early Upper Palaeolithic in Africa. *Science* **314**, 1767–1770 (2006).
28. Patterson, N. *et al.* Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).
29. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
30. Skotte, L., Korneliussen, T. S. & Albrechtsen, A. Estimating individual admixture proportions from next generation sequencing data. *Genetics* **195**, 693–702 (2013).
31. Conard, N. J. A female figurine from the basal Aurignacian of Hohle Fels Cave in southwestern Germany. *Nature* **459**, 248–252 (2009).
32. Straus, L. G. After the deep freeze: confronting “Magdalenian” realities in Cantabrian Spain and beyond. *J. Archaeol. Method Theory* **20**, 236–255 (2013).
33. Weaver, A. J., Saenko, O. A., Clark, P. U. & Mitrovica, J. X. Meltwater pulse 1A from Antarctica as a trigger of the Bolling-Allerød warm interval. *Science* **299**, 1709–1713 (2003).
34. Montoya, C. & Peresani, M. Nouveaux éléments de diachronie dans l'Épigravettien récent des Préalpes de la Vénétie in *D'un monde à l'autre. Les systèmes lithiques pendant le Tardiglaciaire autour de la Méditerranée nord-occidentale* (eds Bracco, J.-P. & Montoya, C.) 123–138 (Mémoire Société Préhistorique Française, 2005).
35. Valentin, B. *Jalons pour une paléohistoire des derniers chasseurs (XIVe-VIe millénaires av. J.-C.)* (Cahiers archéologiques de Paris 1 – 1, Publications de la Sorbonne, 2008).
36. Pala, M. *et al.* Mitochondrial DNA signals of late glacial recolonization of Europe from near eastern refugia. *Am. J. Hum. Genet.* **90**, 915–924 (2012).

**Supplementary Information** is available in the online version of the paper.

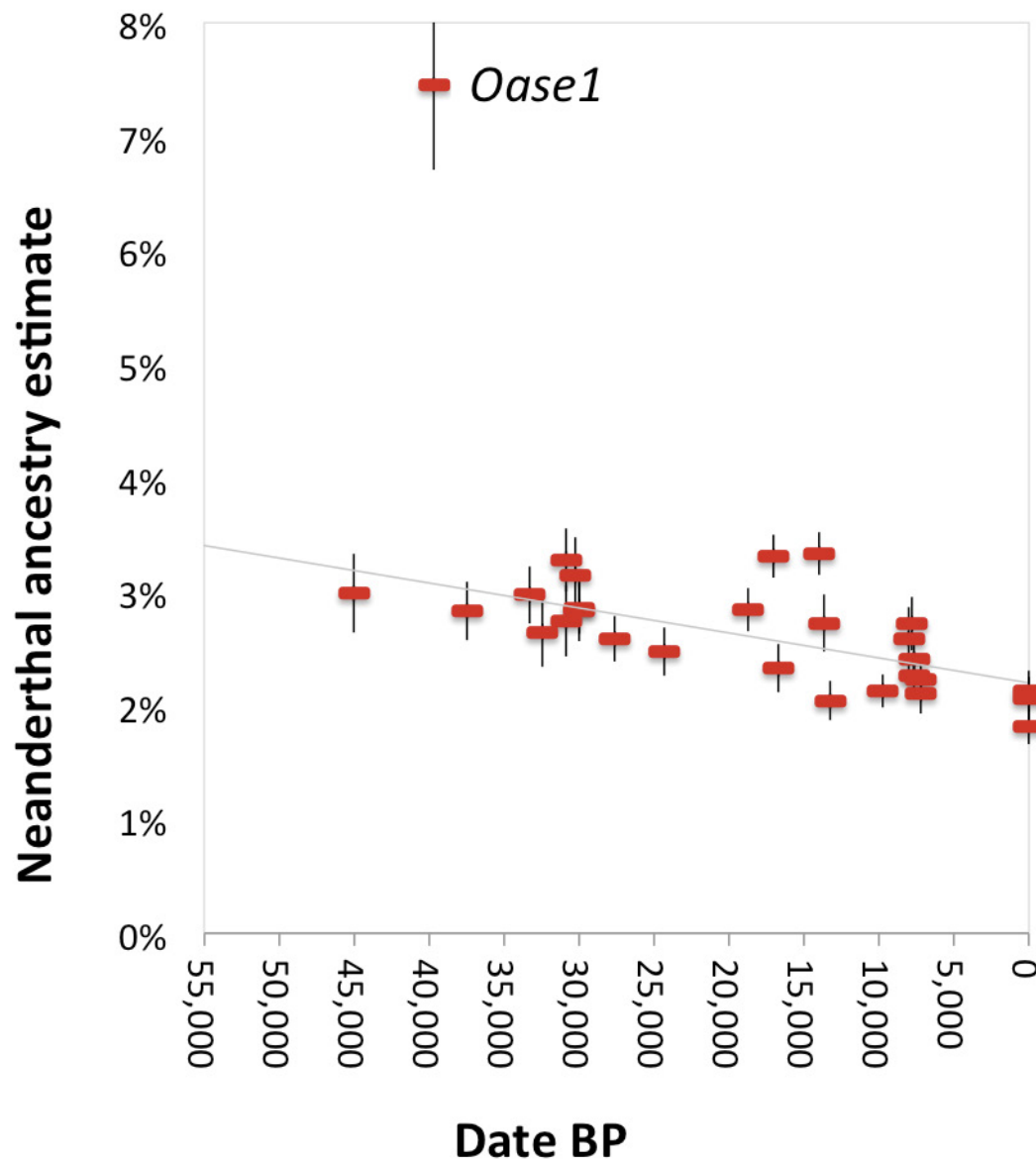
**Acknowledgements** We thank B. Alex, D. Meltzer, P. Moorjani, I. Olalde, S. Sankararaman and B. Viola for comments, K. Stewardson and E. Harney for

sample screening, and F. Hallgren for sharing a radiocarbon date for *Motala12*. The Fig. 1 map is plotted using data available under the Open Database License © OpenStreetMap (<http://www.openstreetmap.org/copyright>). The Goyet project led by H.R. was funded by the Wenner-Gren Foundation (Gr. 7837), the College of Social and Behavioral Sciences of CSUN, the CSUN Competition for Research, Scholarship and Creative Activity Awards, and the RBINS. The excavation of the El Mirón Cave burial, led by L.G.S. and M.R.G.M., was supported by the Gobierno de Cantabria, the L.S.B. Leakey Foundation, the University of New Mexico, the Stone Age Research Fund (J. and R. Auel, principal donors), the town of Ramales de la Victoria and the Universidad de Cantabria. Excavations at Grotta Paglicci were performed by A. Palma di Cesnola in collaboration with the Soprintendenza Archeologia della Puglia (founded by MIUR and local Institutions). Research at Riparo Villabruna was supported by MIBACT and the Veneto Region. Q.F. was funded by the Special Foundation of the President of the Chinese Academy of Sciences (2015–2016), the Bureau of International Cooperation of the Chinese Academy of Sciences, the Chinese Academy of Sciences (XDA05130202), the National Natural Science Foundation of China (L1524016) and the Chinese Academy of Sciences Discipline Development Strategy Project (2015-DX-C-03). D.Fe was supported by an Irish Research Council grant (GOIPG/2013/36). I.M. was supported by a long-term fellowship from the Human Frontier Science Program LT001095/2014-L. P.Sk was supported by the Swedish Research Council (VR 2014-453). S.T. and M.P.R. were funded by the Max Planck Society. C.N.-M. was funded by FWF P-17258, P-19347, P-21660 and P-23612. S.C. and O.T.M. were funded by a ‘Karstives’ Grant PCCE 31/2010 (CNCS-UEFISCDI, Romania). A.P.D., N.D., V.Sla and N.D. were funded by the Russian Science Foundation (project No.14-50-00036). M.A.M. was funded by a Marie Curie Intra-European Fellowship within the 7th European Community Framework Programme (grant number PIEF-GA-2008-219965). M.La and D.C. were funded by grants PRIN 2010-11 and 2010EL8TXP\_003. C.C. and the research about the French Jura sites of Rochedane, Rigney and Ranchot was funded by the Collective Research Program (PCR) (2005-2008). K.H. was supported by the European Research Council (ERC StG 283503) and the Deutsche Forschungsgemeinschaft (DFG INST37/706-1FUGG, DFG FOR2237). D.G.D. was funded by the European Social Fund and Ministry of Science, Research and Arts of Baden-Württemberg. R.P. was funded by ERC starting grant ADNABIOARC (263441). J.Ke was funded by a grant from the Deutsche Forschungsgemeinschaft (SFB1052, project A02). J.Kr was funded by DFG grant KR 4015/1-1, the Baden Württemberg Foundation, and the Max Planck Society. S.P. were funded by the Max Planck Society and the Krekeler Foundation. D.R. was funded by NSF HOMINID grant BCS-1032255, NIH (NIGMS) grant GM100233, and the Howard Hughes Medical Institute.

**Author Contributions** J.Kr, S.P. and D.R. conceived the idea for the study. Q.F., C.P., M.H., W.H., M.Me, V.Slo, R.G.C., A.P.D., N.D., V.Sla, A.T., F.M., B.G., E.V., M.R.G.M., L.G.S., C.N.-M., M.T.-N., S.C., O.T.M., S.B., M.Per, D.Co, M.La, S.R., A.R., F.V., C.T., K.W., D.G., H.R., I.C., D.FI, P.Se, M.A.M., C.C., H.B., N.J.C., K.H., V.M., D.G.D., J.S., D.Ca, R.P., J.Kr, S.P. and D.R. assembled archaeological material. Q.F., C.P., M.H., D.Fe, A.F., W.H., M.Me, A.M., B.N., N.R., V.Slo, S.T., H.B., D.G.D., M.P.R., R.P., J.Kr, S.P. and D.R. performed or supervised wet laboratory work. Q.F., C.P., M.H., M.Pet, S.M., A.P., I.L., M.Li, I.M., S.S., P.Sk, J.Ke, N.P. and D.R. analysed data. Q.F., C.P., M.H., M.Pet, J.Ke, S.P. and D.R. wrote the manuscript and supplements.

**Author Information** The aligned sequences are available through the European Nucleotide Archive under accession number PRJEB13123. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.R. ([reich@genetics.med.harvard.edu](mailto:reich@genetics.med.harvard.edu)).

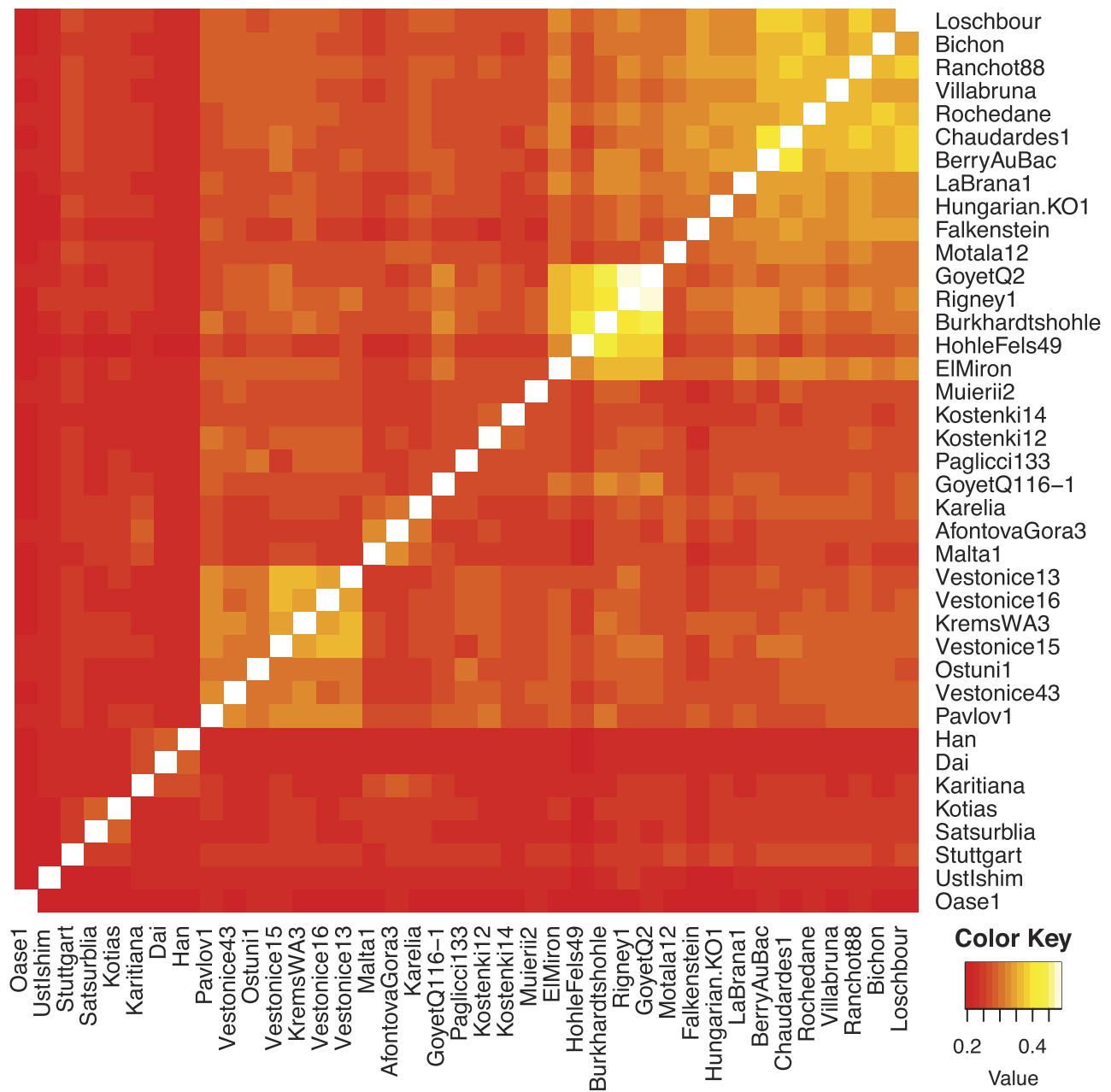
**Reviewer Information** *Nature* thanks C. Lalueza-Fox and the other anonymous reviewer(s) for their contribution to the peer review of this work.



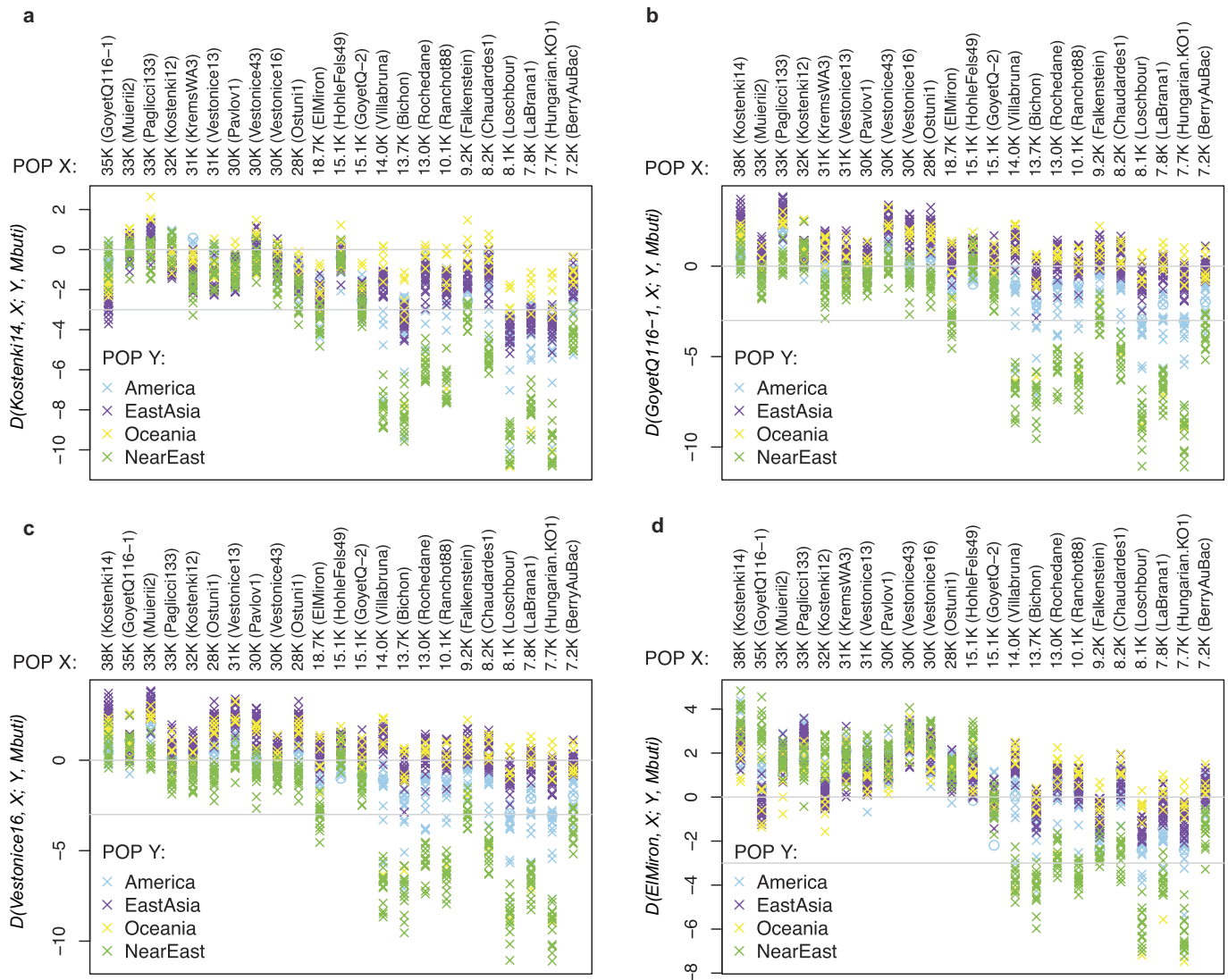
**Extended Data Figure 1 | A decrease in Neanderthal ancestry in the last 45,000 years.** This is similar to Fig. 2, except we use ancestry estimates from rates of alleles matching to Neanderthal rather than  $f_4$ -ratios, as described in Supplementary Information section 3. The least-squares

fit excludes *Oase1* (as an outlier with recent Neanderthal ancestry) and Europeans (known to have reduced Neanderthal ancestry). The regression slope is significantly negative ( $P = 0.00004$ , Extended Data Table 3).





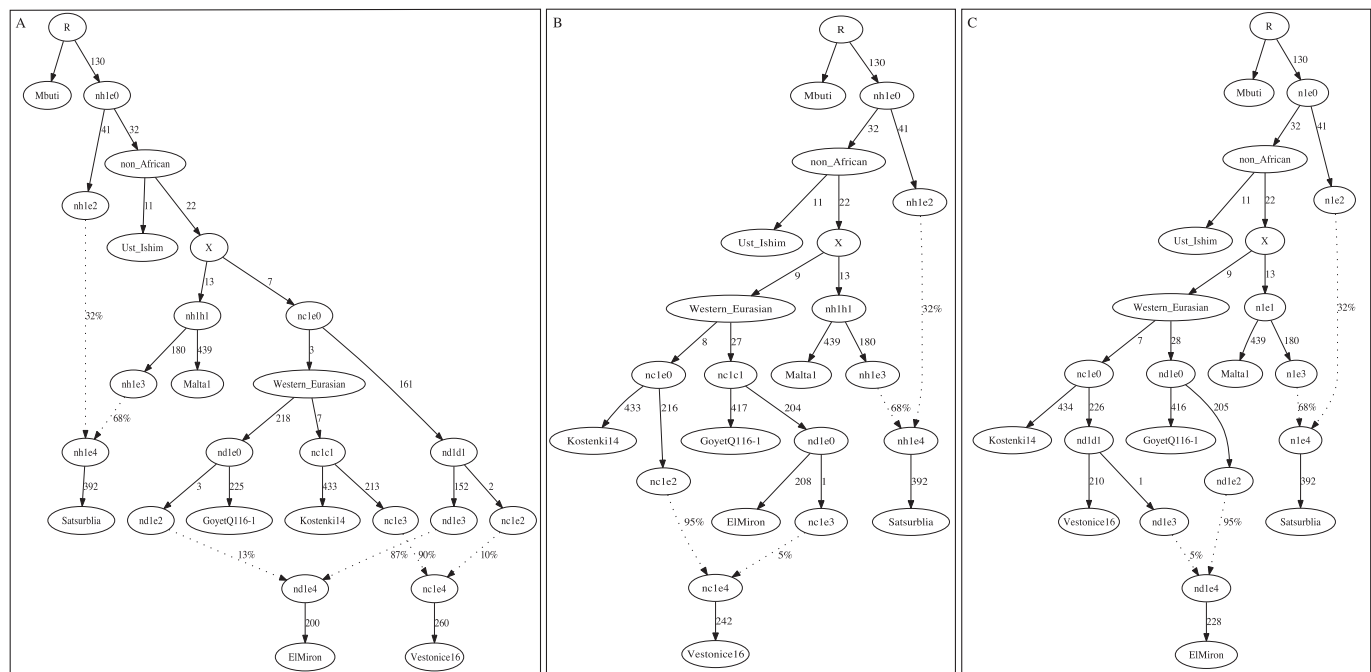
**Extended Data Figure 2 | Heat matrix of pairwise  $f_3(X, Y; Mbuti)$  for selected ancient individuals.** Only individuals with at least 30,000 SNPs covered at least once are analysed.



**Extended Data Figure 3 | Studying how the relatedness of non-European populations to pairs of European hunter-gatherers changes over time.** Statistics were examined of the form  $D(W, X; Y, \text{Mbuti})$ , with the Z-score given on the y axis, where  $W$  is an early European hunter-gatherer,  $X$  is another European hunter-gatherer (in chronological order on the x axis), and  $Y$  is a non-European population (see legend).

**a**,  $W = \text{Kostenki14}$ . **b**,  $W = \text{GoyetQ116-1}$ . **c**,  $W = \text{Vestonice16}$ . **d**,  $W = \text{ElMiron}$ .  $|Z| > 3$  scores are considered statistically significant (horizontal line). The similar Fig. 4b gives absolute  $D$ -statistic values rather than Z-scores (for  $W = \text{Kostenki14}$ ) and uses pooled regions rather than individual populations  $Y$ .





**Extended Data Figure 4 | Three admixture graph models that fit the data for *Satsurblia*, an Upper Palaeolithic individual from the Caucasus.** These models use 127,057 SNPs covered in all populations. Estimated genetic drifts are given along the solid lines in units of  $f_2$ -distance (parts per thousand), and estimated mixture proportions are given along the dotted lines. All three models provide a fit to the

allele frequency correlation data among *Mbuti*, *Ust'-Ishim*, *Kostenki14*, *Vestonice16*, *Malta1*, *ElMiron* and *Satsurblia* to within the limits of our resolution, in the sense that all empirical  $f_2$ -,  $f_3$ - and  $f_4$ -statistics relating the individuals are within three standard errors of the expectation of the model. Models in which *Satsurblia* is treated as unadmixed cannot be fit.

Extended Data Table 1 | The 51 ancient modern humans analysed in this study

Sample Code	Data source	Country	Lat.	Long.	Cal BP 95.4%	Date type (ref.)	Culture	Remain	SNP Panel	Sex	mtDNA haplogroup	Y chrom. haplogroup	Genetic Cluster	Damage restrict	Mean coverage+	SNPs covered
Ustishim	<sup>13</sup>	Russia	57.43	71.10	47,480–42,560	Direct-UF ( <sup>13</sup> )	Unassigned	Femur	Shotgun	M	R	K (xLT)	Unassigned	No	42	2,137,615
Oase1	<sup>3</sup>	Romania	45.12	21.90	41,640–37,580	Direct-UF ( <sup>3</sup> )	Unassigned	Mandible	Shotgun	M	N	F	Unassigned	Yes	0.156	285,076
Kostenki14*	New	Russia	51.23	39.30	38,680–36,260	Direct-UF ( <sup>3</sup> )	Unassigned	Tibia	3.7M	M	U2	C1b	Unassigned	No	16.1	1,774,156
GoyetQ116-1	New	Belgium	50.26	4.28	35,160–34,430	Direct-NotUF ( <sup>3</sup> )	Aurignacian	Humerus	1240k	M	M	C1a	Unassigned	No	1.046	846,983
Muierii2	New	Romania	45.11	23.46	33,760–32,840	Direct-UF ( <sup>3</sup> )	Unassigned	Temporal	3.7M	F	U6		Unassigned	Yes	0.049	98,618
Paglicci33	New	Italy	41.65	15.61	34,580–31,210	Layer ( <sup>3</sup> )	Gravettian	Tooth	1240k	M	U8c	I	Vestonic	No	0.041	82,330
Cioclovina1	New	Romania	45.35	23.84	33,090–31,780	Direct-UF ( <sup>3</sup> )	Unassigned	Cranium	1240k	M	U	CT	Unassigned	Yes	0.006	12,784
Kostenki12	New	Russia	51.23	39.30	32,990–31,840	Layer ( <sup>3</sup> )	Unassigned	Cranium	3.7M	M	U2	CT	Unassigned	No	0.03	61,228
KremsWA3	New	Austria	48.41	15.59	31,250–30,690	Layer ( <sup>3</sup> )	Gravettian	Cranium	1240k	M	U5		Vestonic	No	0.11	203,986
Vestonice13	New	Czech Republic	48.53	16.39	31,070–30,670	Layer ( <sup>3</sup> )	Gravettian	Femur	3.7M	M	U8c	CT(notIJK)	Vestonic	Yes	0.071	139,568
Vestonice15	New	Czech Republic	48.53	16.39	31,070–30,670	Layer ( <sup>3</sup> )	Gravettian	Femur	3.7M	M	U5	BT	Vestonic	Yes	0.015	30,900
Vestonice14	New	Czech Republic	48.53	16.39	31,070–30,670	Layer ( <sup>3</sup> )	Gravettian	Femur	390k	M	U		Vestonic	Yes	0.003	5,677
Pavlov1	New	Czech Republic	48.53	16.39	31,110–29,410	Layer ( <sup>3</sup> )	Gravettian	Femur	3.7M	M	U5	C1a2	Vestonic	Yes	0.028	57,005
Vestonice43	New	Czech Republic	48.53	16.39	30,710–29,310	Layer ( <sup>3</sup> )	Gravettian	Femur	3.7M	M	U	F	Vestonic	Yes	0.087	163,946
Vestonice16	New	Czech Republic	48.53	16.39	30,710–29,310	Layer ( <sup>3</sup> )	Gravettian	Femur	3.7M	M	U5	IJK	Vestonic	No	1.31	945,292
Ostuni2	New	Italy	40.73	17.57	29,310–28,640	Direct-UF (New)	Gravettian	Femur	3.7M	F	U2		Vestonic	Yes	0.008	17,017
GoyetQ53-1	New	Belgium	50.26	4.28	28,230–27,720	Direct-NotUF ( <sup>3</sup> )	Gravettian	Fibula	1240k	F	U2		Vestonic	Yes	0.006	12,567
Paglicci108	New	Italy	41.65	15.61	28,430–27,070	Layer ( <sup>3</sup> )	Gravettian	Phalanx	1240k	F	U2/3'4'7'8'9		Vestonic	Yes	0.002	4,330
Ostuni1	New	Italy	40.73	17.57	27,810–27,430	Direct-UF (New)	Gravettian	Tibia	3.7M	F	M		Vestonic	Yes	0.245	369,313
GoyetQ376-19	New	Belgium	50.26	4.28	27,720–27,310	Direct-NotUF ( <sup>3</sup> )	Gravettian	Humerus	1240k	F	U2		Vestonic	Yes	0.012	25,400
GoyetQ56-16	New	Belgium	50.26	4.28	26,600–26,040	Direct-NotUF ( <sup>3</sup> )	Gravettian	Fibula	1240k	F	U2		Vestonic	Yes	0.005	9,988
Mal'ta	<sup>14</sup>	Russia	52.9	103.5	24,520–24,090	Direct-UF ( <sup>14</sup> )	Unassigned	Humerus	Shotgun	M	U	R	Mal'ta	No	1.174	143,950
ElMirón	New	Spain	43.26	-3.45	18,830–18,610	Direct-UF ( <sup>14</sup> )	Magdalenian	Toe	3.7M	F	U5b		El Mirón	Yes	1.012	797,714
AfontovaGora3	New	Russia	56.05	92.87	16,930–16,490	Layer ( <sup>14</sup> )	Unassigned	Tooth	3.7M	F	R1b		Mal'ta	Yes	0.17	286,355
AfontovaGora2	<sup>14</sup>	Russia	56.05	92.87	16,930–16,490	Direct-UF ( <sup>14</sup> )	Unassigned	Humerus	Shotgun	M	U		Mal'ta	No	0.071	143,751
Rigney1	New	France	47.23	6.10	15,690–15,240	Direct-NotUF ( <sup>14</sup> )	Magdalenian	Mandible	1240k	F	U2/3'4'7'8'9		El Mirón	Yes	0.017	35,600
HohleFels49	New	Germany	48.22	9.45	16,000–14,260	Layer ( <sup>14</sup> )	Magdalenian	Femur	390k	M	U8a	I	El Mirón	Yes	0.033	63,151
GoyetQ-2	New	Belgium	50.26	4.28	15,230–14,780	Direct-NotUF ( <sup>14</sup> )	Magdalenian	Humerus	1240k	M	U8a	H1JK	El Mirón	Yes	0.035	72,263
Brillenohle	New	Germany	48.24	9.46	15,120–14,440	Direct-UF ( <sup>14</sup> )	Magdalenian	Cranium	390k	M	U8a		El Mirón	Yes	0.006	13,459
HohleFels79	New	Germany	48.22	9.45	15,070–14,270	Direct-UF ( <sup>14</sup> )	Magdalenian	Cranium	390k	M	U8a		El Mirón	Yes	0.005	11,211
Burkhardtshöhle	New	Germany	48.32	9.35	15,080–14,150	Direct-NotUF ( <sup>14</sup> )	Magdalenian	Cranium	1240k	M	U8a	I	El Mirón	Yes	0.018	38,376
Villabruna	New	Italy	46.15	12.21	14,180–13,780	Direct-UF ( <sup>14</sup> )	Epigravettian	Femur	3.7M	M	U5b2b	R1b1	Villabruna	No	3.137	1,215,433
Bichon	<sup>2</sup>	Switzerland	47.01	6.79	13,770–13,560	Direct-UF ( <sup>14</sup> )	Azilian	Petrous	Shotgun	M	U5b1h	I2	Villabruna	No	8.119	2,116,782
Satsurbia	<sup>2</sup>	Georgia	42.24	42.92	13,380–13,130	Direct-UF ( <sup>14</sup> )	Epigravettian	Petrous	Shotgun	M	K3	J2	Satsurbia	No	1.195	1,460,368
Rochedane	New	France	47.21	6.45	13,090–12,830	Direct-NotUF ( <sup>14</sup> )	Epipaleolithic	Mandible	1240k	M	U5b2b	I	Villabruna	No	0.131	237,390
Ibousieres39	New	France	44.29	4.46	12,040–11,410	Direct-NotUF ( <sup>14</sup> )	Epipaleolithic	Femur	390k	M	U5b2b		Villabruna	Yes	0.005	9,659
Continenza	New	Italy	41.96	13.54	11,200–10,510	Layer (New)	Mesolithic	Cranium	3.7M	F	U5b1		Villabruna	Yes	0.006	11,717
Ranchol88	New	France	47.91	5.43	10,240–9,930	Direct-NotUF ( <sup>14</sup> )	Mesolithic	Cranium	1240k	F	U5b1		Villabruna	Yes	0.322	414,863
LesCloseaux13	New	France	48.52	2.11	10,240–9,560	Direct-NotUF ( <sup>14</sup> )	Mesolithic	Femur	1240k	F	U5a2		Villabruna	Yes	0.004	8,635
Kotias	<sup>2</sup>	Georgia	42.13	43.12	9,890–9,550	Direct-UF ( <sup>14</sup> )	Mesolithic	Tooth	Shotgun	M	H13c	J	Satsurbia	No	12.157	2,133,968
Falkenstein	New	Germany	48.06	9.04	9,410–8,990	Direct-NotUF ( <sup>14</sup> )	Mesolithic	Fibula	390k	M	U5a2c	F	Villabruna	Yes	0.011	21,977
Karelia	<sup>9</sup>	Russia	61.65	35.65	8,800–7,950	Layer ( <sup>14</sup> )	Mesolithic	Tooth	Shotgun	M	C1g	R1a1	Unassigned	No	1.952	1,754,410
Bockstein	New	Germany	48.33	10.09	8,370–8,160	Layer ( <sup>14</sup> )	Mesolithic	Tooth	390k	F	U5b1d1		Villabruna	Yes	0.011	21,977
Ofnet	New	Germany	48.49	10.27	8,430–8,060	Layer ( <sup>14</sup> )	Mesolithic	Tooth	390k	F	U5b1d1		Villabruna	Yes	0.003	6,263
Chaudardes1	New	France	49.24	3.46	8,360–8,050	Direct-NotUF ( <sup>14</sup> )	Mesolithic	Tibia	1240k	M	U5b1b	I	Villabruna	Yes	0.046	92,657
Loschbour	<sup>15</sup>	Luxembourg	49.70	6.24	8,160–7,940	Direct-UF ( <sup>14</sup> )	Mesolithic	Tooth	Shotgun	M	U5b1a	I2a1b	Villabruna	No	20	2,091,584
LaBran1	<sup>16</sup>	Spain	42.93	-5.35	7,940–7,690	Direct-UF ( <sup>14</sup> )	Mesolithic	Tooth	Shotgun	M	U5b2c1	C1a2	Villabruna	No	3.338	1,884,745
Hungarian.KO1	<sup>17</sup>	Hungarian	47.93	21.20	7,730–7,590	Direct-UF ( <sup>14</sup> )	Neolithic	Petrous	Shotgun	M	R3	I2a	Villabruna	No	1.1	1,410,303
Motala12	<sup>15</sup>	Sweden	58.54	15.05	7,670–7,580	Direct-UF (New)	Mesolithic	Tooth	Shotgun	M	U2c1	I2a1b*	Unassigned	No	2.185	1,874,519
BerryAubac	New	France	49.24	3.54	7,320–7,170	Direct-NotUF ( <sup>14</sup> )	Mesolithic	Radius	1240k	M	U5b1a	I	Villabruna	No	0.027	54,690
Stuttgart	<sup>15</sup>	Germany	48.78	9.18	7,260–7,020	Direct-UF (New)	Early Neolithic	Tooth	Shotgun	F	T2c1d1		Unassigned	No	19	2,078,724

Refs 37–57 are cited in this table. All dates are obtained as described in Supplementary Information section 1. When an individual has a direct date from the same skeleton it is marked 'direct' followed by a hyphen to indicate whether the date is obtained by ultrafiltration ('UF') or without ('NotUF'). If the date is from the archaeological layer, the date type is marked as 'layer'. All the dates are calibrated using IntCal13 (ref. 58) and the OxCal4.2 program<sup>59</sup>.

\*Kostenki14 is represented in most analyses by our newly reported 16.1× capture data, but key analyses were repeated on the previously reported 2.8× shotgun data<sup>4</sup>.

+Mean coverage is computed on the 3.7 million SNP targets.

37. Rougier, H. *et al.* Peștera cu Oase 2 and the cranial morphology of early modern Europeans. *Proc. Natl Acad. Sci. USA* **104**, 1165–1170 (2007).
38. Marom, A., McCullagh, J. S. O., Higham, T. F. G., Sinitsyn, A. A. & Hedges, R. E. M. Single amino acid radiocarbon dating of Upper Paleolithic modern humans. *Proc. Natl Acad. Sci. USA* **109**, 6878–6881 (2012).
39. Soficaru, A., Dobos, A. & Trinkaus, E. Early modern humans from the Peștera Muierii, Baia de Fier, Romania. *Proc. Natl Acad. Sci. USA* **103**, 17196–17201 (2006).
40. Palma di Cesnola, A. *Paglicci. L'Aurignaziano e il Gravettiano antico* (Claudio Grenzi, 2004).
41. Soficaru, A., Petrea, C., Dobos, A. & Trinkaus, E. The human cranium from the Peștera Cioclovina Uscată, Romania — context, age, taphonomy, morphology, and paleopathology. *Curr. Anthropol.* **48**, 611–619 (2007).
42. Sinitsyn, A. A. Les sépultures de Kostenki: chronologie, attribution culturelle, rite funéraire in *La spiritualité* (Otte M. (ed.)), Proceedings of UISPP conference, Liège, ERAUL **106**, 237–244 (2004).
43. Simon, U., Haendel, M., Einwoegerer, T. & Neugebauer-Maresch, C. The archaeological record of the Gravettian open air site Krems-Wachtberg. *Quat. Int.* **351**, 5–13 (2014).
44. Trinkaus, E. & Svoboda, J. *Early Modern Human Evolution in Central Europe: The People of Dolní Věstonice and Pavlov*, Vol. 12 (Oxford Univ. Press, 2006).
45. Azzi, C. M., Bigliocca, L. & Piovani, F. Florence radiocarbon dates II. *Radiocarbon* **16**, 10–14 (1974).
46. Cupillard, C. *et al.* Changes in ecosystems, climate and societies in the Jura Mountains between 40 and 8 ka cal BP. *Quat. Int.* **378**, 40–72 (2015).
47. Housley, R. A., Gamble, C. S., Street, M. & Pettitt, P. B. Radiocarbon evidence for the Lateglacial human recolonisation of Northern Europe. *Proc. Prehist. Soc.* **63**, 25–54 (1997).
48. Benazzi, S. *et al.* Early dispersal of modern humans in Europe and implications for Neanderthal behaviour. *Nature* **479**, 525–528 10.1038/nature10617 (2011).
49. Simon, U. Die Burkhardtshöhle — eine Magdalénienstation am Nordrand der Schwäbischen Alb, Magisterarbeit (1993).
50. Vercellotti, G., Alciati, G., Richards, M. P. & Formicola, V. The Late Upper Paleolithic skeleton Villabruna 1 (Italy): a source of data on biology and behavior of a 14,000 year-old hunter. *J. Anthropol. Sci.* **86**, 143–163 (2008).
51. Drucker, D. G., Bridault, A., Cupillard, C., Hujic, A. & Bocherens, H. Evolution of habitat and environment of red deer (*Cervus elaphus*) during the Late-glacial and early Holocene in eastern France (French Jura and the western Alps) using multi-isotope analysis ( $\delta^{13}\text{C}$ ,  $\delta^{15}\text{N}$ ,  $\delta^{18}\text{O}$ ,  $\delta^{34}\text{S}$ ) of archaeological remains. *Quat. Int.* **245**, 268–278 (2011).
52. Valentin, F. *et al.* Découvertes récentes d'inhumations et d'une incinération datées du Mésolithique en Île-de-France. *Revue Archéologique d'Île-de-France* **1**, 21–42 (2008).
53. Bramanti, B. *et al.* Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. *Science* **326**, 137–140 (2009).
54. Price, T. D. & Jacobs, K. Olenii Ostrov — first radiocarbon dates from a major mesolithic cemetery Karelia, USSR. *Antiquity* **64**, 849–853 (1990).
55. Wehrberger, K. "Der Streit ward definitiv beendet..." Eine mesolithische Bestattung aus der Bocksteinhöhle im Lonetal, Alb-Donau-Kreis. Zur Erinnerung an Ludwig Bürger (1844–1898). *Archäologisches Korrespondenzblatt* **30**, 15–31 (2000).
56. Orschiedt, J. Manipulationen an menschlichen Skelettresten. Taphonomische Prozesse, Sekundärbestattungen oder Kannibalismus. Dissertation. Urgeschichtliche Materialhefte 13 (1999).
57. Sánchez-Quinto, F. *et al.* Genomic affinities of two 7,000-year-old Iberian hunter-gatherers. *Curr. Biol.* **22**, 1494–1499 (2012).
58. Reimer, P. J. *et al.* IntCal13 and marine13 radiocarbon age calibration curves 0–50,000 years cal BP. *Radiocarbon* **55**, 1869–1887 (2013).
59. Ramsey, C. B. & Lee, S. Recent and planned developments of the program OxCal. *Radiocarbon* **55**, 720–730 (2013).



Extended Data Table 2 | Estimated proportion of Neanderthal ancestry

		$f_T$ -ratios					Archaic Ancestry Informative SNPs						
Sample Code	Age BP	SNPs	Est.	95% CI			SNPs	Est.	95% CI		Increase in Neanderthal ancestry with B	S.E.	
UstIshim	45,020	2,137,615	4.4%	3.6%	-	5.3%	778,774	3.0%	2.3%	-	3.7%	-0.9%	1.3%
Oase1	39,610	285,076	9.9%	8.4%	-	11.4%	59,854	7.5%	6.0%	-	8.9%	2.5%	1.8%
Kostenki14	37,470	1,774,156	3.6%	2.7%	-	4.4%	632,748	2.8%	2.3%	-	3.3%	-1.0%	1.0%
GoyetQ116-1	34,795	846,983	3.4%	2.4%	-	4.3%							
Muierii2	33,300	98,618	5.2%	3.0%	-	7.4%	22,189	3.0%	2.5%	-	3.5%	0.6%	1.1%
Paglicci133	32,895	82,330	4.1%	2.1%	-	6.0%							
Ciocolovina1	32,435	12,784	4.1%	-1.1%	-	9.3%							
Kostenki12	32,415	61,228	1.9%	-0.7%	-	4.4%	13,385	2.6%	2.1%	-	3.2%	1.7%	1.5%
KremsWA3	30,970	203,986	3.9%	2.6%	-	5.2%				-			
Vestonice13	30,870	139,568	4.6%	2.6%	-	6.5%	35,983	3.3%	2.7%	-	3.8%	0.3%	1.3%
Vestonice15	30,870	30,900	4.3%	0.6%	-	7.9%	5,855	2.7%	2.1%	-	3.4%	-1.5%	1.3%
Vestonice14	30,870	5,677	2.6%	-5.9%	-	11.0%							
Pavlov1	30,260	57,005	4.4%	1.6%	-	7.1%	9,327	3.1%	2.5%	-	3.8%	0.7%	1.2%
Vestonice43	30,010	163,946	6.9%	5.2%	-	8.5%	38,749	2.9%	2.4%	-	3.3%	0.9%	0.9%
Vestonice16	30,010	945,292	4.1%	3.1%	-	5.1%	268,157	2.8%	2.3%	-	3.3%	-0.1%	1.0%
Ostuni2	28,975	17,017	1.6%	-3.2%	-	6.3%	2,746	2.3%	1.4%	-	3.1%	1.3%	1.6%
GoyetQ53-1	27,975	12,567	4.8%	-0.7%	-	10.3%							
Paglicci108	27,750	4,330	3.4%	-6.0%	-	12.7%							
Ostuni1	27,620	369,313	4.2%	3.0%	-	5.4%	88,449	2.6%	2.2%	-	3.0%	0.1%	0.9%
GoyetQ376-19	27,515	25,400	6.5%	2.7%	-	10.2%							
GoyetQ56-16	26,320	9,988	3.6%	-1.9%	-	9.1%							
Malta1	24,305	1,439,501	2.9%	1.9%	-	3.8%	437,187	2.5%	2.1%	-	2.9%	1.0%	0.8%
ElMiron	18,720	797,714	3.6%	2.6%	-	4.5%	250,071	2.8%	2.5%	-	3.2%	0.6%	0.9%
AfontovaGora3	16,710	286,355	3.0%	1.8%	-	4.2%	96,237	3.3%	2.9%	-	3.7%	-1.5%	1.0%
AfontovaGora2	16,710	143,751	2.2%	0.4%	-	4.0%	37,280	2.3%	1.9%	-	2.7%	-0.3%	0.9%
Rigney1	15,465	35,600	0.8%	-2.6%	-	4.2%							
HohleFels49	15,130	63,151	2.3%	-0.6%	-	5.2%							
GoyetQ-2	15,005	72,263	1.7%	-0.6%	-	4.0%							
Brillenhohle	14780	13,459	2.5%	-3.0%	-	8.1%							
HohleFels79	14,670	11,211	1.7%	-5.1%	-	8.5%							
Burkhardtshohle	14,615	38,376	1.7%	-1.6%	-	5.0%							
Villabruna	13,980	1,215,433	2.7%	1.8%	-	3.5%	425,148	3.3%	3.0%	-	3.7%	1.1%	0.9%
Bichon	13,665	2,116,782	2.9%	1.9%	-	3.8%	769,422	2.7%	2.2%	-	3.2%	0.7%	1.3%
Satsurbli	13,255	1,460,368	1.5%	0.6%	-	2.4%	542,561	2.0%	1.7%	-	2.4%	0.9%	0.6%
Rochedane	12,960	237,390	1.9%	0.5%	-	3.3%							
Iboussieres39	11,725	9,659	6.4%	-0.8%	-	13.7%							
Continenza	10,855	11,717	4.1%	-1.4%	-	9.6%	1,733	2.9%	1.8%	-	4.0%	-10.6%	4.4%
Ranchot88	10,085	414,863	2.9%	1.8%	-	4.0%							
LesClosaux13	9,900	8,635	-	-9.7%	-	3.8%							
Kotias	9,720	2,133,968	1.8%	1.0%	-	2.7%	779,146	2.1%	1.8%	-	2.4%	0.7%	0.5%
Falkenstein	9,200	64,428	4.8%	1.7%	-	7.8%							
Karelia	8,375	1,754,410	1.9%	1.1%	-	2.7%	582,444	2.2%	1.9%	-	2.6%	-0.2%	0.7%
Bockstein	8,265	21,977	5.7%	1.0%	-	10.5%							
Ofnet	8,245	6,263	9.8%	1.4%	-	18.1%							
Chaudardes1	8,205	92,657	1.9%	-0.2%	-	3.9%							
Loschbour	8,050	2,091,584	2.5%	1.6%	-	3.3%	774,139	2.6%	2.0%	-	3.1%	2.7%	1.7%
LaBranal	7,815	1,884,745	1.9%	1.1%	-	2.8%	642,231	2.7%	2.3%	-	3.2%	0.4%	0.8%
Hungarian.KO1	7,660	1,410,303	2.1%	1.2%	-	3.0%	439,408	2.4%	2.0%	-	2.8%	-0.1%	1.2%
Motala12	7,625	1,874,519	2.5%	1.6%	-	3.3%	655,685	2.3%	1.9%	-	2.7%	-0.1%	0.7%
BerryAuBac	7,245	54,690	2.5%	-0.2%	-	5.1%							
Stuttgart	7,140	2,078,724	1.9%	1.1%	-	2.7%	767,813	2.1%	1.8%	-	2.5%	0.0%	0.7%
Dai	0	2,144,502	1.4%	0.7%	-	2.1%	782,066	1.8%	1.5%	-	2.1%	1.4%	0.4%
Han	0	2,144,502	1.8%	1.1%	-	2.5%	782,164	2.1%	1.8%	-	2.5%	1.9%	0.7%
English	0	2,144,502	1.5%	0.8%	-	2.2%							
French	0	2,144,502	1.5%	0.9%	-	2.1%	782,386	1.7%	1.4%	-	1.9%	1.4%	0.6%
Sardinian	0	2,144,502	1.2%	0.6%	-	1.9%	782,351	1.7%	1.4%	-	2.0%	0.7%	0.5%
Karitiana	0						782,037	2.1%	1.7%	-	2.4%	1.5%	1.0%

**Extended Data Table 3 | Significant correlation of Neanderthal ancestry estimate with specimen age**

Subset of samples	N	P-value for date correlation	Decrease in ancestry per 10,000 years	Estimate of Neanderthal ancestry at different time points			
				0 years ago (present)	50,000 years ago	55,000 years ago	60,000 years ago
<b><i>f<sub>I</sub></i>-ratio estimates</b>							
Core Set 1 (all ancient samples (except <i>Oase1</i> ) + <i>Han</i> + <i>Dai</i> )	57	$5 \times 10^{-22}$	0.48-0.73%	1.1-2.2%	4.0-5.4%	4.3-5.7%	4.5-6.0%
Subset of Core Set 1 (<32kya)	50	$2 \times 10^{-15}$	0.59-0.98%	0.9-2.1%	4.5-6.4%	4.8-6.9%	5.1-7.4%
Subset of Core Set 1 (>32kya or <25kya)	44	$4 \times 10^{-18}$	0.44-0.69%	1.0-2.2%	3.7-5.2%	4.0-5.5%	4.2-5.8%
Subset of Core Set 1 (>25kya or <14kya)	47	$5 \times 10^{-21}$	0.48-0.73%	1.0-2.2%	3.9-5.3%	4.2-5.7%	4.5-6.0%
Subset of Core Set 1 (>14kya or present day)	37	$2 \times 10^{-18}$	0.47-0.74%	1.1-2.4%	4.1-5.5%	4.3-5.8%	4.6-6.2%
Subset of Core Set 1 (only ancient samples)	50	$4 \times 10^{-15}$	0.46-0.76%	1.0-2.3%	4.0-5.4%	4.3-5.8%	4.5-6.1%
Subset of Core Set 1 (individuals with >200,000 SNPs)	28	$4 \times 10^{-19}$	0.46-0.71%	1.1-2.3%	3.9-5.3%	4.2-5.7%	4.4-6.0%
Modification of Core Set 1 (replace East Asians with Europeans)	58	$2 \times 10^{-23}$	0.49-0.73%	1.1-2.3%	4.0-5.4%	4.3-5.8%	4.6-6.1%
All ancient samples including <i>Oase1</i> + <i>Han</i> + <i>Dai</i>	58	$8 \times 10^{-29}$	0.57-0.81%	1.0-2.2%	4.3-5.7%	4.7-6.1%	5.0-6.5%
All ancient samples	51	$1 \times 10^{-20}$	0.57-0.86%	0.9-2.2%	4.4-5.8%	4.7-6.2%	5.0-6.6%
All ancient samples except <i>Oase1</i> or <i>UstIshim</i>	49	$8 \times 10^{-12}$	0.45-0.81%	1.0-2.3%	4.0-5.6%	4.2-6.0%	4.5-6.4%
<b>Ancestry informative SNPs</b>							
Core Set 2 (all ancient samples (except <i>Oase1</i> ) + <i>Han</i> + <i>Dai</i> + <i>Karitiana</i> )	29	$4 \times 10^{-11}$	0.21-0.39%	1.8-2.3%	3.1-4.0%	3.2-4.2%	3.3-4.3%
Subset of Core Set 2 (no <i>Han</i> , <i>Dai</i> , <i>Karitiana</i> , <i>Stuttgart</i> )	25	$1 \times 10^{-4}$	0.11-0.36%	1.8-2.5%	2.9-3.8%	3.0-4.0%	3.0-4.1%
Subset of Core Set 2 (no <i>Han</i> , <i>Dai</i> , <i>Karitiana</i> , <i>Stuttgart</i> , <i>UstIshim</i> )	24	$2 \times 10^{-4}$	0.11-0.37%	1.8-2.5%	2.9-3.8%	2.9-4.0%	3.0-4.2%

'Core set 1' used for the *f<sub>I</sub>*-ratio analyses, refers to 50 ancient individuals (removing *Oase1* as an outlier) along with 7 east Asians (*Dai* and *Han*). 'Core set 2' used for the analyses of Neanderthal ancestry informative SNPs, refers to 26 ancient individuals (removing *Oase1*, *Han*, *Dai* and *Karitiana*).



**Extended Data Table 4 | Sex determination for newly reported individuals**

Sample	Target	Type	N <sub>auto</sub>	N <sub>x</sub>	N <sub>y</sub>	N <sub>x</sub> /N <sub>auto</sub>	N <sub>y</sub> /N <sub>auto</sub>	X-rate	Y-rate	Sex
	1240k or 2.2M		1151240	49711	32681	0.0432	0.0284			
	390k		388745	1819	2242	0.0047	0.0058			
Kostenki14	2.2M	all	29633405	395534	262846	0.0133	0.0089	0.309	0.312	M
GoyetQ116-1	1240k	all	2122620	36391	22256	0.0171	0.0105	0.397	0.369	M
Cioclovina1	1240k	Damage	11521	184	125	0.0160	0.0108	0.370	0.382	M
Kostenki12	2.2M	Subset	63908	856	504	0.0134	0.0079	0.310	0.278	M
Muierii2	2.2M	Damage	81165	2177	8	0.0268	0.0001	0.621	0.003	F
Vestonice13	2.2M	Damage	119094	1578	1059	0.0133	0.0089	0.307	0.313	M
Vestonice15	2.2M	Damage	28762	338	227	0.0118	0.0079	0.272	0.278	M
Vestonice14	390k	Damage	4846	8	11	0.0017	0.0023	0.353	0.394	M
Vestonice43	2.2M	Damage	136933	1826	1204	0.0133	0.0088	0.309	0.310	M
Pavlov1	2.2M	Damage	54429	631	404	0.0116	0.0074	0.268	0.261	M
Vestonice16	2.2M	Subset	2433741	30463	20976	0.0125	0.0086	0.290	0.304	M
KremsWA3	1240k	all	235069	4119	2661	0.0175	0.0113	0.406	0.399	M
Ostuni2	2.2M	Damage	15749	138	1	0.0088	0.0001	0.203	0.002	F
Ostuni1	2.2M	Damage	427199	10868	47	0.0254	0.0001	0.589	0.004	F
Paglicci108	1240k	Damage	3883	124	2	0.0319	0.0005	0.740	0.018	F
GoyetQ53-1	1240k	Damage	10771	311	4	0.0289	0.0004	0.669	0.013	F
GoyetQ376-19	1240k	Damage	20052	680	10	0.0339	0.0005	0.785	0.018	F
GoyetQ56-16	1240k	Damage	8702	304	7	0.0349	0.0008	0.809	0.028	F
Paglicci133	1240k	Subset	81092	1641	983	0.0202	0.0121	0.469	0.427	M
ElMiron	2.2M	Damage	1765696	40647	196	0.0230	0.0001	0.533	0.004	F
HohleFels79	390k	Damage	10188	28	22	0.0027	0.0022	0.587	0.374	M
AfontovaGora3	2.2M	Damage	291798	8705	37	0.0298	0.0001	0.691	0.004	F
HohleFels49	390k	Damage	61051	113	111	0.0019	0.0018	0.396	0.315	M
Rigney1	1240k	Damage	32797	1131	9	0.0345	0.0003	0.799	0.010	F
GoyetQ-2	1240k	Damage	65563	1123	706	0.0171	0.0108	0.397	0.379	M
Brillenhohle	390k	Damage	12603	22	22	0.0017	0.0017	0.373	0.303	M
Burkhardtshohle	1240k	Damage	34207	563	407	0.0165	0.0119	0.381	0.419	M
Villabruna	2.2M	Subset	5505838	72055	52110	0.0131	0.0095	0.303	0.333	M
Rochedane	1240k	Subset	256325	4780	2830	0.0186	0.0110	0.432	0.389	M
Continenza	2.2M	Damage	10647	208	2	0.0195	0.0002	0.452	0.007	F
Ibousseries39	390k	Damage	8246	12	22	0.0015	0.0027	0.311	0.463	M
Ranchot88	1240k	Damage	594962	18520	119	0.0311	0.0002	0.721	0.007	F
LesCloseaux13	1240k	Damage	7326	275	2	0.0375	0.0003	0.869	0.010	F
Falkenstein	390k	Damage	58970	113	102	0.0019	0.0017	0.410	0.300	M
Bockstein	390k	Damage	20214	62	0	0.0031	0.0000	0.655	0.000	F
Ofnet	390k	Damage	5294	13	1	0.0025	0.0002	0.525	0.033	F
Chaudardes1	1240k	Damage	84052	1429	865	0.0170	0.0103	0.394	0.363	M
BerryAuBac	1240k	All	49670	902	554	0.0182	0.0112	0.421	0.393	M

\*We restrict analysis to the 1240k target set for study of the 2.2M capture datasets.

Y-rate is the ratio of  $N_y/N_{\text{auto}}$  divided by the same quantity for the genome-wide target set. Female sex (F) is inferred as Y-rate < 0.05 and male sex (M) as Y-rate > 0.

Extended Data Table 5 | Allele counts at SNPs affected by selection in individuals with &gt;1-fold coverage

SNP		<i>LCT</i>	<i>SLC45A2</i>	<i>SLC24A5</i>	<i>EDAR</i>	<i>HERC2</i>
		rs4988235	rs16891982	rs1426654	rs3827760	rs12913832
Ancestral		G	C	G	A	A
Derived		A	G	A	G	G
UstIshim	Coverage	31	46	52	42	50
	Derived allele frequency	0%	0%	2%	0%	0%
Kostenki14	Coverage	140	113	6	45	52
	Derived allele frequency	0%	2%	17%	0%	0%
GoyetQ116-1	Coverage	8	6	0	9	1
	Derived allele frequency	0%	0%	n/a	0%	0%
Vestonice16	Coverage	13	18	0	4	5
	Derived allele frequency	0%	6%		0%	0%
Malta1	Coverage	1	0	2	2	2
	Derived allele frequency	0%		0%	0%	0%
ElMiron	Coverage	2	10	0	7	5
	Derived allele frequency	0%	0%		0%	0%
Villabruna	Coverage	17	52	5	19	10
	Derived allele frequency	0%	0%	0%	0%	100%
Bichon	Coverage	11	4	25	16	9
	Derived allele frequency	0%	0%	0%	0%	33%
Satsurbliia	Coverage	1	2	4	1	4
	Derived allele frequency	0%	0%	100%	0%	50%
Kotias	Coverage	16	22	13	20	15
	Derived allele frequency	0%	0%	100%	0%	0%
Loschbour	Coverage	19	18	20	17	21
	Derived allele frequency	0%	0%	0%	0%	100%
LaBranal	Coverage	8	6	2	11	3
	Derived allele frequency	12%	0%	0%	0%	100%
Hungarian.KO1	Coverage	1	2	2	1	2
	Derived allele frequency	0%	0%	50%	0%	100%
Motala12	Coverage	2	0	3	3	1
	Derived allele frequency	0%		0%	33%	100%
Karelia	Coverage	1	9	4	0	1
	Derived allele frequency	0%	67%	0%		0%
Stuttgart	Coverage	25	21	15	29	21
	Derived allele frequency	0%	0%	100%	0%	0%

rs4988235 is responsible for lactase persistence in Europe<sup>60,61</sup>. The SNPs at *SLC24A5* and *SLC45A2* are responsible for light skin pigmentation. The SNP at *EDAR*<sup>62,63</sup> affects tooth morphology and hair thickness. The SNP at *HERC2* (refs 64, 65) is the primary determinant of light eye colour in present-day Europeans. We present the fraction of fragments overlapping each SNP that are derived; the observation of a low rate of derived alleles does not prove that the individual carried the allele, and instead may reflect sequencing error or ancient DNA damage. Sites highlighted in light grey were judged (based on the derived allele count) likely to be heterozygous for the derived allele, and dark grey sites are likely to be homozygous.

60. Enattah, N. S. *et al.* Identification of a variant associated with adult-type hypolactasia. *Nature Genet.* **30**, 233–237 (2002).
61. Bersaglieri, T. *et al.* Genetic signatures of strong recent positive selection at the lactase gene. *Am. J. Hum. Genet.* **74**, 1111–1120 (2004).
62. Fujimoto, A. *et al.* A scan for genetic determinants of human hair morphology: *EDAR* is associated with Asian hair thickness. *Hum. Mol. Genet.* **17**, 835–843 (2008).
63. Kimura, R. *et al.* A common variation in *EDAR* is a genetic determinant of shovel-shaped incisors. *Am. J. Hum. Genet.* **85**, 528–535 (2009).
64. Sturm, R. A. *et al.* A single SNP in an evolutionary conserved region within intron 86 of the *HERC2* gene determines human blue-brown eye color. *Am. J. Hum. Genet.* **82**, 424–431 (2008).
65. Eiberg, H. *et al.* Blue eye color in humans may be caused by a perfectly associated founder mutation in a regulatory element located within the *HERC2* gene inhibiting *OCA2* expression. *Hum. Genet.* **123**, 177–187 (2008).

**Extended Data Table 6 | All European hunter-gatherers beginning with *Kostenki14* share genetic drift with present-day Europeans**

Test	SNPs used	D-value	Z score
Ust'-Ishim	2,050,358	0.003	6.6
Oase1	278,785	0.005	10.6
Kostenki14	1,676,253	-0.002	-5.5
Muierii2	95,787	-0.004	-6.3
GoyetQ116-1	811,756	-0.004	-8.0
Kostenki12	59,850	-0.004	-5.1
Paglicci133	79,624	-0.004	-5.5
Vestonice13	136,598	-0.004	-7.1
Vestonice15	30,252	-0.006	-6.4
Vestonice16	914,141	-0.004	-9.1
Pavlov1	55,835	-0.005	-6.3
Vestonice43	160,463	-0.004	-6.9
KremsWA3	229,187	-0.005	-10.2
Ostuni1	360,347	-0.004	-8.6
Malta1	1,401,718	-0.005	-11.3
ElMiron	777,654	-0.007	-14.7
AfontovaGora2	141,073	-0.007	-13.6
AfontovaGora3	707,617	-0.006	-13.6
HohleFels49	62,816	-0.004	-5.2
Rigney1	34,445	-0.006	-6.1
GoyetQ-2	70,210	-0.006	-8.8
Burkhardtshohle	37,234	-0.006	-6.2
Villabruna	1,170,777	-0.010	-24.7
Bichon	2,034,069	-0.009	-23.6
Satsurbliia	1,419,824	-0.005	-13.1
Rochedane	229,806	-0.011	-20.8
Ranchot88	402,274	-0.010	-21.8
Kotias	2,047,856	-0.006	-15.8
Falkenstein	64,043	-0.008	-11.6
Chaudardes1	90,047	-0.011	-16.0
Loschbour	2,037,082	-0.011	-25.4
LaBrana1	1,824,307	-0.009	-23.0
Motala12	1,816,201	-0.009	-23.8
Hungarian.KO1	1,372,801	-0.010	-26.5
Karelia	1,701,664	-0.009	-21.9
Stuttgart	2,023,939	-0.009	-23.9
BerryAuBac	53,028	-0.011	-14.0

The statistic  $D(Han, Test; French, Mbuti)$  was computed measuring whether present-day French share more alleles with *Han* or with a *Test* population (restricting to ancient individuals with at least 30,000 SNPs covered at least once). Present-day Europeans share significantly more genetic drift with European hunter-gatherers from *Kostenki14* onward than they do with *Han*. Thus, by the date of *Kostenki14*, there was already west Eurasian-specific genetic drift.



# Midbrain circuits for defensive behaviour

Philip Tovote<sup>1\*</sup>, Maria Soledad Esposito<sup>1,2\*</sup>, Paolo Botta<sup>1†</sup>, Fabrice Chaudun<sup>3</sup>, Jonathan P. Fadok<sup>1</sup>, Milica Markovic<sup>1</sup>, Steffen B. E. Wolff<sup>1†</sup>, Charu Ramakrishnan<sup>4</sup>, Lief Fenno<sup>4</sup>, Karl Deisseroth<sup>4</sup>, Cyril Herry<sup>3</sup>, Silvia Arber<sup>1,2</sup> & Andreas Lüthi<sup>1</sup>

**Survival in threatening situations depends on the selection and rapid execution of an appropriate active or passive defensive response, yet the underlying brain circuitry is not understood. Here we use circuit-based optogenetic, *in vivo* and *in vitro* electrophysiological, and neuroanatomical tracing methods to define midbrain periaqueductal grey circuits for specific defensive behaviours. We identify an inhibitory pathway from the central nucleus of the amygdala to the ventrolateral periaqueductal grey that produces freezing by disinhibition of ventrolateral periaqueductal grey excitatory outputs to pre-motor targets in the magnocellular nucleus of the medulla. In addition, we provide evidence for anatomical and functional interaction of this freezing pathway with long-range and local circuits mediating flight. Our data define the neuronal circuitry underlying the execution of freezing, an evolutionarily conserved defensive behaviour, which is expressed by many species including fish, rodents and primates. In humans, dysregulation of this ‘survival circuit’ has been implicated in anxiety-related disorders.**

Threatening situations, such as the presence of a predator or exposure to stimuli predicting imminent or perceived danger, evoke an evolutionarily conserved brain state, fear, which triggers defensive behaviours to avoid or reduce potential harm<sup>1–3</sup>. A long-standing question in fear and anxiety research has been how brain circuits generate various forms of defensive behaviours, which have been used as a read-out for normal fear and maladaptive anxiety<sup>2,4–6</sup>. In rodents, depending on threat imminence<sup>7</sup> and contextual factors such as the existence of escape routes, defensive behaviours range from risk assessment<sup>8</sup> and freezing<sup>9,10</sup> to flight and defensive attack<sup>2,11</sup>. These behaviours can be rapidly switched to adequately adapt to fluctuating threat levels or contextual challenges<sup>11,12</sup>. On the basis of electrical stimulation, lesion and pharmacological studies, the midbrain periaqueductal grey region (PAG) has been proposed to present an essential part of the circuitry that elicits freezing and flight in response to threat<sup>12–22</sup>. However, PAG circuit mechanisms underlying expression of defensive behaviours remain poorly understood. This includes a lack of knowledge about the functional roles of different PAG neuron types, their connectivity and regulation for expression of defensive behaviours. The PAG receives inputs from key forebrain regions involved in regulation of defensive behaviour, such as the central nucleus of the amygdala (CEA)<sup>23,24</sup>, the hypothalamus<sup>25–27</sup> and medial prefrontal cortex<sup>28</sup>, but little is known about their specific PAG cellular targets. In addition, the functional roles of long-range inputs to PAG, neuronal subpopulations within PAG subregions and intra-PAG microcircuitry, as well as outputs from PAG, in the expression of active and passive defensive behaviours is poorly understood. Using optogenetic manipulations of specific cell types, single-unit recordings and rabies-mediated neuroanatomical tracings, we here define a pathway from the CEA to the ventrolateral PAG (vPAG) that mediates freezing by disinhibition of vPAG outputs to pre-motor targets in the magnocellular nucleus (Mc) of the medulla. Furthermore, we provide evidence for anatomical and functional interaction of this ‘freezing pathway’ with circuits mediating flight.

## Freezing is mediated by glutamatergic vPAG neurons

To determine cellular diversity in the vPAG that could be associated with a distinct behavioural phenotype, we used an optogenetic

approach to specifically manipulate the activity of excitatory glutamatergic neurons, one of the main cell classes in the PAG. We targeted glutamatergic neurons expressing vesicular glutamate transporter 2 (vGluT2<sup>+</sup>) by local injection of adeno-associated viruses (AAV) delivering a construct that contained a Cre-dependent channel-rhodopsin-2 (ChR2) coupled to an mCherry tag into the vPAG of *Vglut2-ires-Cre* mice (Fig. 1a, b and Extended Data Fig. 1a, b). Mice injected with AAVs containing a fluorescent tag only served as controls. We first optically manipulated cellular activity in naive mice under low-fear conditions (that is, at low freezing levels), during exposure to a novel context. Strikingly, light-activation of vGluT2<sup>+</sup> neurons of the vPAG reliably triggered strong freezing behaviour during the ‘light on’ period (Fig. 1c and Supplementary Video 1), which was reflected by a marked decrease in behavioural activity (Fig. 1d). To define the endogenous function of vPAG glutamatergic neurons, we next used viral-vector-mediated, Cre-dependent expression of archaerhodopsin (Arch) to optically inhibit these neurons (Fig. 1b and Extended Data Fig. 1c). Under low-fear conditions in naive mice, we did not observe an effect on freezing or behavioural activity (Fig. 1e, f). To investigate the necessity of vPAG glutamatergic neurons for conditioned freezing behaviour, we subjected mice to auditory fear conditioning, followed by next-day re-exposure to an aversively conditioned tone stimulus (CS) or to the context previously paired with mild electrical foot-shocks. We found that optical inhibition of vPAG vGluT2<sup>+</sup> neurons during the CS blocked tone-induced freezing (Fig. 1g and Supplementary Video 2). Similarly, freezing in the conditioning context was markedly reduced during optical inhibition of vPAG vGluT2<sup>+</sup> neurons (Fig. 1h). These results demonstrate a role of vPAG glutamatergic neurons in mediating conditioned freezing responses.

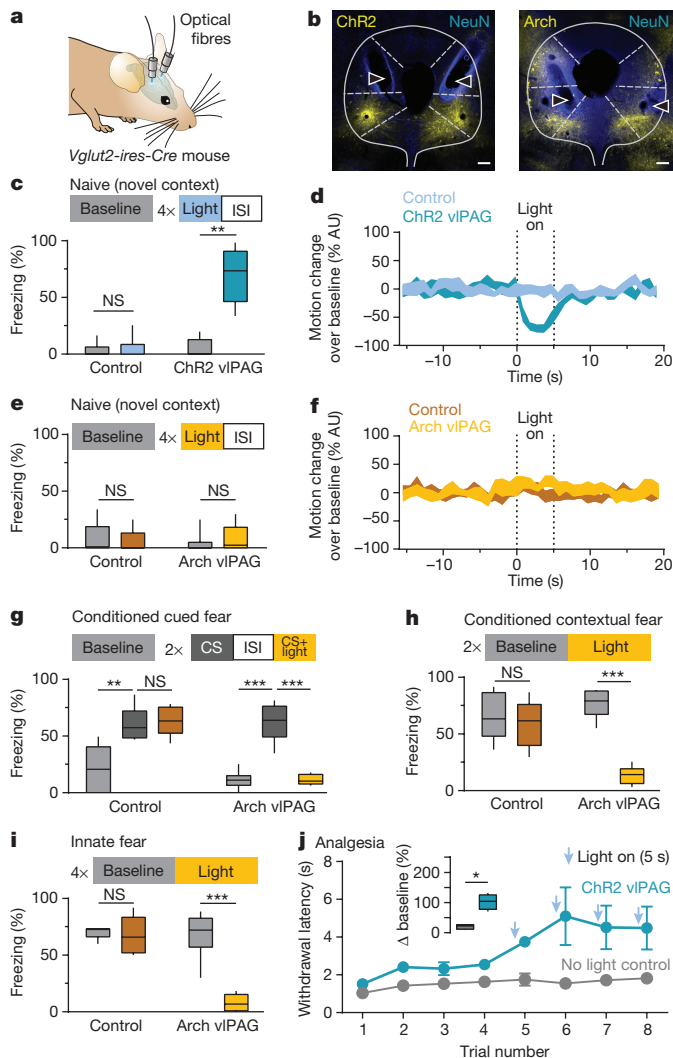
To determine whether these neurons serve a more general role in mediating freezing, we next tested the necessity of these neurons for producing freezing to an innate threat. Mice exhibit strong fear responses when exposed to large moving objects, probably because these resemble visual features of a natural predator<sup>22,29</sup>. We therefore exposed mice in an open-field arena to a remote controlled toy

<sup>1</sup>Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland. <sup>2</sup>Biozentrum, Department of Cell Biology, University of Basel, 4056 Basel, Switzerland.

<sup>3</sup>INSERM, Neurocentre Magendie, U862, 146 Rue Léo-Saignat, Bordeaux 33077, France. <sup>4</sup>Stanford University, 318 Campus Drive West, Clark Center W080, Stanford, California 94305, USA.

<sup>†</sup>Present addresses: Champalimaud Centre for the Unknown, Avenida de Brasília, 1400-038 Lisbon, Portugal (P.B.); Center for Brain Science, Harvard University, Cambridge, Massachusetts 02138, USA (S.B.E.W.).

\*These authors contributed equally to this work.



**Figure 1 | Glutamatergic vPAG neurons drive defensive responses.** **a**, Optogenetics in freely moving mice. **b**, Expression patterns of ChR2 (left) and Arch (right) within vGluT2<sup>+</sup> vPAG neurons (triangles, fibre tracts; scale bars, 200  $\mu$ m). **c, d**, Light activation of glutamatergic vPAG neurons triggered freezing behaviour ( $n = 10$  ChR2,  $n = 12$  control, two-tailed Wilcoxon signed-rank test). **e, f**, Inhibition of vPAG glutamatergic neurons had no effect on freezing or behavioural activity in naive mice ( $n = 7$  Arch,  $n = 12$  control, two-tailed Wilcoxon signed-rank test). **g–i**, Inhibition of vPAG glutamatergic neurons diminished CS-induced freezing ( $n = 6$  per group,  $1 \times 3$  analysis of variance (ANOVA),  $F_{(2,10)} = 37.72$ ,  $P < 0.001$ , Tukey's post-hoc test), contextual freezing ( $n = 6$  Arch,  $n = 8$  control, paired two-tailed Student's  $t$ -test) and innate freezing responses ( $n = 6$  Arch,  $n = 5$  control, paired two-tailed Student's  $t$ -test). **j**, Light activation of vPAG glutamatergic neurons induced analgesia ( $n = 6$  ChR2,  $n = 10$  control, unpaired two-tailed Student's  $t$ -test). Box-whisker plots indicate median, interquartile range and 5th–95th percentiles of the distribution. Motion plots depict s.e.m. range. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ .

snake (Extended Data Fig. 1d). We found that while mice exhibited strong freezing responses in the presence of the fear stimulus, this reaction was dramatically reduced by yellow-light-mediated inhibition of vPAG vGluT2<sup>+</sup> neurons (Fig. 1i and Extended Data Fig. 1e). The reduction in freezing was attributable to optical inhibition and not to changes in threat imminence as measured by spatial distance between mouse and snake (Extended Data Fig. 1f, h). Taken together, these findings demonstrate that activation of vPAG glutamatergic neurons is necessary for both learned and innate freezing, and that it can generate freezing in the absence of threat.

Because the PAG is involved in processing ascending and descending pain information from the periphery<sup>30–32</sup>, it is conceivable that the observed behavioural responses were related to enhanced light-evoked pain perception. However, when we tested nociception using a tail immersion test, we found that optical activation of vGluT2<sup>+</sup> vPAG neurons had a marked analgesic effect (Fig. 1j). These experiments identify a cellular substrate in the vPAG for analgesia, an important part of the defensive response to threat.

### A disinhibitory pathway from CEA to vPAG

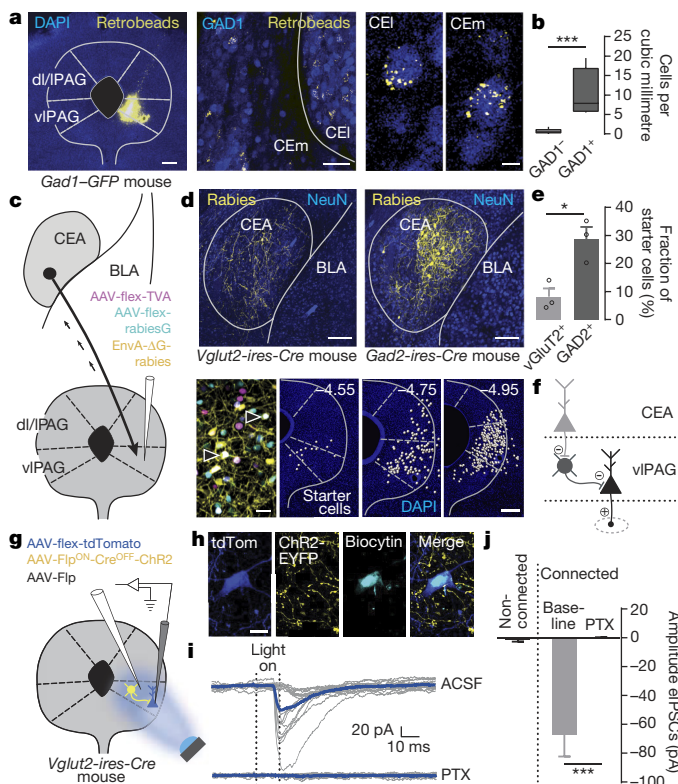
To address the question of how vPAG vGluT2<sup>+</sup> neurons are regulated, we first aimed to characterize inputs to the vPAG from the CEA, because of their suggested major roles in the expression of freezing<sup>5,12,33,34</sup>. To identify the CEA neurons contacting vPAG, we injected retrogradely transported red fluorescent latex beads into the vPAG of a reporter mouse strain, in which GABAergic ( $\gamma$ -aminobutyric-acid-releasing) cells express enhanced green fluorescent protein (*Gad1-eGFP*; Fig. 2a). Retrogradely transported beads were found throughout the CEA (Fig. 2a). Quantification of overlap between beads and GFP<sup>+</sup> neurons (Fig. 2b) showed that the CEA sends a GABAergic projection to vPAG, consistent with previous reports<sup>24</sup>. Freezing is elicited through enhanced CEA output<sup>33,35</sup>, but since these neurons are GABAergic, this would result in enhanced inhibition of their vPAG targets. However, our results from optogenetic manipulation of freezing behaviour suggest that freezing is associated with increased activity of vPAG glutamatergic neurons. Thus, the most parsimonious explanation consistent with these observations would involve a local vPAG disinhibitory circuit mechanism that could convert an inhibitory input from CEA into enhanced output of vPAG glutamatergic neurons.

To test this hypothesis, we traced monosynaptic connections of CEA cells onto either glutamatergic or GABAergic cells in the vPAG (Fig. 2c) using Cre-dependent, cell-type-specific infection with pseudotyped EnvA-G-deleted rabies virus (EnvA- $\Delta$ G-rabies)<sup>36</sup> in *Vglut2-ires-Cre* or *Gad2-ires-Cre* mouse lines. These experiments revealed that CEA projections preferentially target vPAG GABAergic cells (Fig. 2d, e). Complementary evidence for a disinhibitory CEA–vPAG circuit was provided by whole-cell patch-clamp recordings of vPAG GABAergic cells during optical activation of CEA terminals in acute brainstem slices containing PAG (Extended Data Fig. 2a). This experiment confirmed the existence of such inhibitory connections to vPAG GABAergic neurons (Extended Data Fig. 2b–f and Fig. 2f). We also probed the existence of functional connections between local GABAergic and glutamatergic cells in the vPAG. Glutamatergic cells were visualized by viral-vector-mediated, Cre-dependent expression of tdTomato in the vPAG of *Vglut2-ires-Cre* mice (Fig. 2g). Using a double-conditional viral approach<sup>37</sup>, ChR2 was introduced into vPAG of non-vGluT2<sup>+</sup> cells. We recorded from identified vGluT2<sup>+</sup> neurons in acute brain slices during optical activation of local, non-vGluT2<sup>+</sup> neurons (Fig. 2h). In 50% of all recorded vGluT2<sup>+</sup> cells, we observed optically evoked inhibitory postsynaptic potentials (eIPSCs; average latency = 5.5 ms) (Fig. 2i) that were completely blocked by application of picrotoxin, a GABA<sub>A</sub> receptor antagonist (Fig. 2j). Taken together, our data from slice recordings and tracing experiments suggest that freezing is caused by a disinhibitory process within vPAG that involves CEA-mediated inhibition of local GABAergic neurons resulting in enhanced activity of glutamatergic vPAG neurons (Fig. 2f).

### Local vPAG GABAergic neurons control freezing

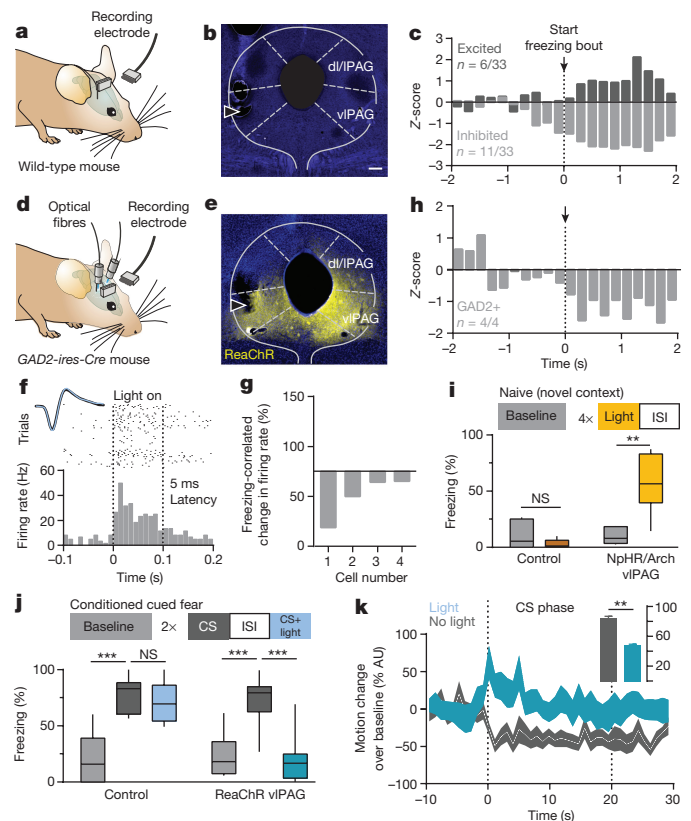
To investigate whether *in vivo* neuronal correlates of freezing in the vPAG would be consistent with a disinhibitory circuit mechanism, we next performed single-unit recordings within vPAG in freely moving mice (Fig. 3a). Mice with chronically implanted recording electrodes (Fig. 3b and Extended Data Fig. 3a) were fear conditioned and freezing responses were evoked during a fear retrieval session 24 h





**Figure 2 | A disinhibitory pathway from CEA to vIPAG.** **a**, Fluorescent latex beads in vIPAG (left, scale bar, 200  $\mu$ m). Beads were found in medial (CEm) and lateral (CEl) CEA cells (middle, scale bar, 40  $\mu$ m; zoom-in shown in two right panels, scale bar, 5  $\mu$ m). **b**, Overlap between beads and GAD1<sup>+</sup> or GAD1<sup>-</sup> neurons ( $n = 2$  mice, two-tailed Mann-Whitney test). **c**, Cell-type-specific monosynaptic rabies tracing strategy. **d**, Rabies-labelled cells within CEA of *Vglut2-Cre* or *GAD2-Cre* mice (top panels, scale bar, 100  $\mu$ m). Starter cells coexpressing TVA-GFP, rabiesG-V5 and EnvA- $\Delta$ G-mCherry-rabies (black triangles, bottom left panel, scale bar, 25  $\mu$ m). Example of starter cells (white dots, three bottom right panels; numbers indicate distance from bregma; scale bar, 200  $\mu$ m). **e**, Normalized number of rabies cells in CEA ( $n = 3$  *Vglut2-Cre* mice,  $n = 3$  *GAD2-Cre* mice, unpaired two-tailed Student's *t*-test). **f**, Schematic model of a disinhibitory pathway from CEA to vIPAG. **g**, **h**, Expression of ChR2 in non-glutamatergic neurons and tdTomato in glutamatergic neurons for targeted whole-cell patch-clamp recordings (scale bar, 10  $\mu$ m). **i**, **j**, Light-evoked IPSCs (observed in 50% of all glutamatergic neurons tested) were blocked by PTX application ( $n = 12$  cells from six slices of four mice, two-tailed  $1 \times 3$  ANOVA,  $F_{(2,15)} = 18.88$ ,  $P < 0.0001$ , Sidak's post-hoc test). Box-whisker plots indicate median, interquartile range and 5th–95th percentiles of the distribution; bar plots indicate mean  $\pm$  s.e.m. \* $P < 0.05$ , \*\*\* $P < 0.001$ .

later<sup>38</sup>. Principal component analysis of single-unit activity in relation to freezing behaviour revealed that while one population of neurons (18%) was activated during freezing bouts, another population (33%) was inhibited (Fig. 3c). Consequently, we examined whether glutamatergic and GABAergic neurons could contribute to these differential activity patterns. To address this question, we performed recordings from optogenetically identified<sup>39</sup> GAD2<sup>+</sup> and vGluT2<sup>+</sup> neurons within vIPAG (Fig. 3d–f and Extended Data Fig. 3b). In agreement with a disinhibitory vIPAG freezing circuit, light-identified GAD2<sup>+</sup> neurons ( $n = 4$ ; Fig. 3g) exhibited a relatively high median baseline firing rate (8.3 Hz), and all showed lower firing rates during freezing compared with non-freezing periods (Fig. 3g, h). Recordings from identified vGluT2<sup>+</sup> units ( $n = 6$ , median baseline firing rate = 3.0 Hz; Extended Data Fig. 3c–e) during freezing periods revealed a more heterogeneous picture, suggesting the existence of multiple subpopulations of glutamatergic neurons in vIPAG. Thus, the *in vivo* correlates



**Figure 3 | GABAergic vIPAG neurons control freezing.** **a**, Single-unit recordings in the vIPAG of freely moving wild-type mice. **b**, Example of a recording site (triangle; scale bar, 200  $\mu$ m). **c**, vIPAG neuronal populations showing increased or decreased activity during freezing ( $n = 8$  mice; bin size, 10 ms). **d**, Optical identification of GAD2<sup>+</sup> single units in the vIPAG. **e**, Example of ChR2 expression and recording site (triangle; scale bar, 200  $\mu$ m). **f**, Identified GAD2<sup>+</sup> neuron activated by light with short latency (5 ms; bin size, 10 ms). Inset: mean spontaneous and light-evoked spike waveform. **g**, **h**, GAD2<sup>+</sup> neurons exhibited reduced firing rates during freezing (four cells from three mice). **i**, Optical inhibition of GAD2<sup>+</sup> neurons induced freezing in naive mice ( $n = 6$  per group, two-tailed paired Student's *t*-test). **j**, **k**, Optical activation of GAD2<sup>+</sup> neurons impaired CS-evoked freezing ( $n = 12$  per group, two-tailed Friedman test,  $P < 0.0001$ , Dunn's multiple comparison test), and shifted CS-induced fear responses towards active behaviour ( $n = 12$  per group, two-tailed paired Student's *t*-test). Box-whisker plots indicate median, interquartile range, and 5th–95th percentiles of the distribution; bars indicate mean  $\pm$  s.e.m. Motion plots depict s.e.m. range. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .

of freezing are consistent with a disinhibitory circuit design leading to the activation of a subpopulation of glutamatergic vIPAG neurons during freezing.

An important prediction of this model is that manipulating the activity of GABAergic vIPAG neurons should affect freezing. In line with this interpretation, optogenetic inhibition of vIPAG GAD2<sup>+</sup> neurons resulted in markedly enhanced freezing levels in naive animals (Fig. 3i and Extended Data Fig. 3f). Importantly, activation of GAD2<sup>+</sup> neurons reduced freezing in response to a conditioned tone (Fig. 3j and Extended Data Fig. 3g). Moreover, this manipulation not only reduced CS-induced freezing but also resulted in transiently enhanced locomotor activity, resembling flight responses (Fig. 3k). Freezing was also reduced by optical activation of GAD2<sup>+</sup> neurons during re-exposure to the conditioning context (Extended Data Fig. 3h), as well as in the presence of an unconditioned threatening stimulus (Extended Data Fig. 3i), while it had no effect on freezing in low-fear conditions (a novel context; Extended Data Fig. 3j). Taken together, these data are fully compatible with a circuit organization wherein inhibition of



local GABAergic vPAG neurons leads to activation of glutamatergic neurons that is both necessary and sufficient to induce freezing behaviour.

### vPAG output to Mc drives freezing

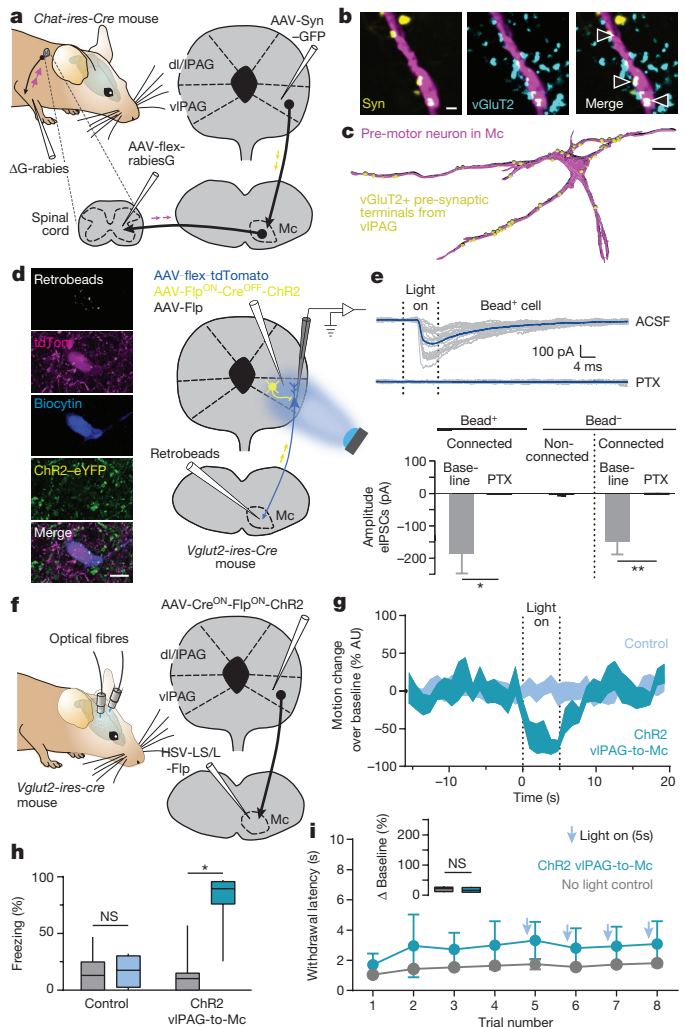
We next sought to identify the output pathway mediating the freezing response. Analysis of axonal projections and synaptic boutons of vGluT2<sup>+</sup> vPAG neurons (Extended Data Fig. 4a) showed that medullary regions previously implicated in motor control, such as the Mc<sup>40,41</sup>, were a major synaptic target. To determine whether vGluT2<sup>+</sup> vPAG neurons target pre-motor neurons in the Mc directly, we performed monosynaptic rabies tracing from forelimb motor neurons in adult *Chat-ires-Cre* mice (Fig. 4a). We visualized glutamatergic vPAG terminals by unconditional viral expression of the presynaptic marker synaptophysin-GFP (Syn-GFP; Extended Data Fig. 4b) with synaptic co-localization of vGluT2 (Fig. 4b). Analysis of presynaptic inputs to forelimb pre-motor Mc neurons revealed that they were directly contacted by glutamatergic vPAG neurons (Fig. 4c and Extended Data Fig. 4c–e).

We then asked whether vPAG Mc-projecting glutamatergic neurons receive local inhibitory input. We first used a monosynaptic intersectional rabies tracing approach to specifically label presynaptic neurons in the vPAG projecting onto glutamatergic neurons which target Mc (Extended Data Fig. 4f). We found that one-third of the local presynaptic neurons were GAD1<sup>+</sup> (Extended Data Fig. 4g, h), and thus putative sources of GABAergic inhibition onto glutamatergic vPAG output cells. Furthermore, we probed the existence of functional connections between local GABAergic and Mc-projecting glutamatergic vPAG cells by combining the approach described in Fig. 2g with an injection of retrogradely transported latex beads into the Mc (Fig. 4d). All Mc-projecting vGluT2<sup>+</sup> neurons recorded in whole-cell patch-clamp showed eIPSCs induced by light activation of vPAG non-vGluT2<sup>+</sup> cells, which were blocked by picrotoxin application (Fig. 4e). These findings provide evidence that excitatory vPAG output to the Mc is under local GABAergic control and suggest that this pathway could be part of the disinhibitory circuit underlying freezing.

To establish functional relevance of this projection for the freezing response, we next used an intersectional optogenetic approach to specifically manipulate activity of the glutamatergic vPAG-to-Mc projection. We injected into the Mc of *Vglut2-ires-Cre* mice retrogradely trafficked herpes simplex virus (HSV), which Cre-dependently expresses flipase (Flp)<sup>42</sup>. This allowed us to selectively introduce ChR2 into vPAG glutamatergic neurons projecting to Mc (vPAG-to-Mc) on the basis of their co-expression of both Cre and Flp using double conditional AAVs (Fig. 4f and Extended Data Fig. 4i–k). Optical activation of glutamatergic vPAG-to-Mc neurons resulted in instantaneous and strong freezing behaviour as reflected by decreased behavioural activity (Fig. 4g, h, Extended Data Fig. 4l and Supplementary Video 3). Interestingly, and in contrast to the anti-nociceptive effect elicited upon projection-unspecific activation of vGluT2<sup>+</sup> neurons in the vPAG (Fig. 1j), no analgesia was observed after stimulation of the glutamatergic vPAG-to-Mc neurons (Fig. 4i). Together, these findings demonstrate that vPAG-to-Mc glutamatergic projection neurons are specifically involved in the expression of freezing behaviour and suggest that analgesia and freezing could be mediated by distinct subpopulations of vPAG glutamatergic neurons.

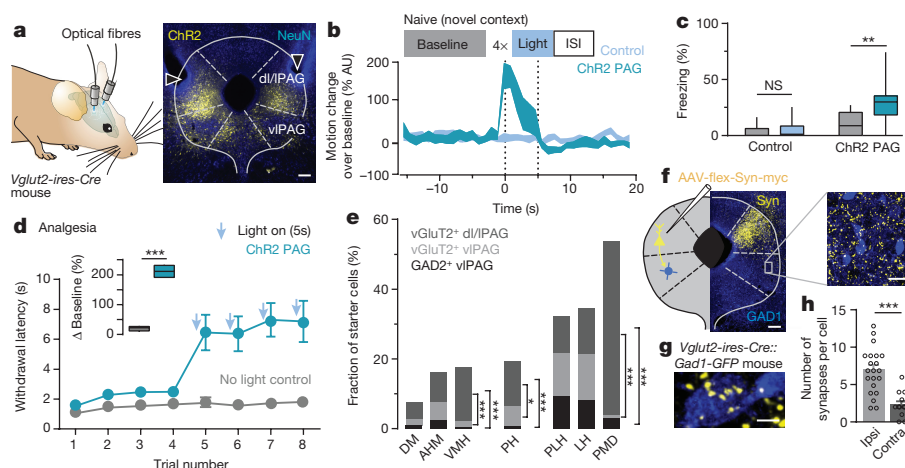
### Interactions between freezing and flight pathways

Our data show that activation of vPAG vGluT2<sup>+</sup> neurons induces freezing (Fig. 1). However, in mice with viral expression extending to dl/PAG, we observed radically different, active defensive behavioural responses through optical activation of vGluT2<sup>+</sup> neurons (Fig. 5a and Extended Data Fig. 5a). These consisted of strong light-induced locomotor activity (Fig. 5b), amounting to marked flight responses



**Figure 4 | Glutamatergic output to the Mc drives freezing.** **a**, Strategy to assess vPAG input onto Mc pre-motor neurons. **b**, vPAG glutamatergic synapses contacting Mc pre-motor neurons were identified by co-staining of Syn-GFP and vGluT2 (triangles; scale bar, 2  $\mu$ m). **c**, Three-dimensional reconstruction of a Mc pre-motor neuron contacted by vPAG vGluT2<sup>+</sup> synapses (yellow circles; scale bar, 10  $\mu$ m). **d**, Intersectional approach to investigate connectivity of vPAG GABAergic cells with Mc-projecting glutamatergic neurons. **e**, Light activation of non-glutamatergic fibres evoked IPSCs in all bead<sup>+</sup> Mc-projecting glutamatergic neurons tested (nine out of nine cells, one example cell shown), and six out of nine glutamatergic, bead<sup>−</sup> neurons. All eIPSCs were blocked by PTX application (four slices of four mice; bead<sup>+</sup>: two-tailed Wilcoxon signed-rank test; bead<sup>−</sup>: 1  $\times$  3 ANOVA  $F_{(2,10)} = 10.28$ ,  $P < 0.01$ , Sidak's post-hoc test). **f**, Strategy to express ChR2 selectively in glutamatergic vPAG neurons projecting to Mc. **g**, **h**, Light activation of glutamatergic vPAG neurons projecting to Mc in naive mice resulted in strong freezing ( $n = 7$  ChR2,  $n = 10$  control, two-tailed Wilcoxon signed-rank test). **i**, Light activation of the glutamatergic vPAG-to-Mc projection had no effect on nociception ( $n = 5$  ChR2,  $n = 11$  control, unpaired two-tailed Student's  $t$ -test). Box-whisker plots indicate median, interquartile range and 5th–95th percentiles of the distribution; bar plots indicate mean  $\pm$  s.e.m. Motion plots depict s.e.m. range. \* $P < 0.05$ ; \*\* $P < 0.01$ .

in some cases. Nonetheless, bouts of forward locomotor activity were often interrupted by short periods of freezing, which resulted, on average, in intermediate levels of freezing (Fig. 5c and Extended Data Fig. 5b). Broad activation of vGluT2<sup>+</sup> neurons within PAG also resulted in strong analgesia (Fig. 5d). Given the observation of alternating active and passive defensive responses, this strongly suggests intricate interactions between flight and freezing circuits and raises



**Figure 5 | Interactions between PAG freezing and flight pathways.** **a**, Optogenetic manipulation of vGluT2<sup>+</sup> neurons in dl/IPAG and vlPAG (triangles indicate fibre tracts; scale bar, 200  $\mu$ m). **b**, **c**, Light activation of PAG glutamatergic neurons in naive mice resulted in flight responses and intermediate freezing levels ( $n = 12$  ChR2,  $n = 12$  control, two-tailed Wilcoxon signed-rank test). **d**, Light activation of PAG glutamatergic neurons resulted in strong analgesia ( $n = 9$  ChR2,  $n = 10$  control, unpaired two-tailed Student's *t*-test). **e**, Hypothalamic presynaptic inputs onto vGluT2<sup>+</sup> ( $n = 3$  mice for each vlPAG and dl/IPAG) and GAD2<sup>+</sup> ( $n = 4$  mice) neurons within vlPAG or dl/IPAG ( $1 \times 3$  ANOVA, Tukey's post-hoc test for each analysed region). Quantification reveals differential input of hypothalamic subregions to vGluT2<sup>+</sup> ( $n = 3$  mice for each vlPAG and dl/IPAG) and GAD2<sup>+</sup> ( $n = 4$  mice) neurons within vlPAG or dl/IPAG. Ventromedial hypothalamic nucleus (VMH) ( $1 \times 3$  ANOVA,  $F_{(2,7)} = 40.22$ ,  $P < 0.0001$ , Tukey's post-hoc test), posterior hypothalamic nucleus (PH) ( $1 \times 3$  ANOVA,  $F_{(2,7)} = 21.01$ ,  $P < 0.05$ , Tukey's post-hoc test) and

premamillary nucleus (PMD) ( $1 \times 3$  ANOVA,  $F_{(2,7)} = 287$ ,  $P < 0.0001$ , Tukey's post-hoc test) preferentially target dl/IPAG vGluT2<sup>+</sup> neurons, whereas dorsomedial hypothalamic nucleus (DM) ( $1 \times 3$  ANOVA,  $F_{(2,7)} = 4.082$ ,  $P > 0.05$ ), anterior hypothalamus, medial part (AHM) ( $1 \times 3$  ANOVA,  $F_{(2,7)} = 3.294$ ,  $P > 0.05$ ), peduncular part of lateral hypothalamus (PLH) ( $1 \times 3$  ANOVA,  $F_{(2,7)} = 0.153$ ,  $P > 0.05$ ) and lateral hypothalamic area (LH) ( $1 \times 3$  ANOVA,  $F_{(2,7)} = 0.948$ ,  $P > 0.05$ ) showed no preference. **f**, Expression of Syn-myc in glutamatergic terminals of dl/IPAG projections to GABAergic neurons within vlPAG (left scale bar, 200  $\mu$ m; right scale bar, 20  $\mu$ m). **g**, High-resolution image of a GAD1<sup>+</sup> vlPAG neuron with dl/IPAG input (scale bar, 5  $\mu$ m). **h**, Quantification of dl/IPAG vGluT2<sup>+</sup> synapses onto vlPAG GAD1<sup>+</sup> ( $n = 21$  ipsilateral cells from four mice,  $n = 13$  contralateral cells from two mice, unpaired two-tailed Student's *t*-test). Box-whisker plots indicate median, interquartile range and 5th–95th percentiles of the distribution; bar plots indicate mean  $\pm$  s.e.m.; motion plots depict s.e.m. \* $P < 0.05$ ; \*\*\* $P < 0.001$ .

the question of how such interactions are implemented within PAG circuitry.

We therefore examined whether distinct presynaptic inputs differentially connect onto specific PAG neuronal subpopulations (Extended Data Fig. 5c). Monosynaptic, cell-specific rabies tracing revealed high selectivity in the hypothalamic projections to defined PAG neuronal subpopulations (Fig. 5e and Extended Data Fig. 5d–f). It is conceivable that hypothalamic input to both dl/IPAG and vlPAG glutamatergic neurons promotes a range of defensive behaviours<sup>26,27</sup>. In turn, activity of the disinhibitory pathway originating in the CEA might bias the behavioural response towards freezing instead of flight. We thus hypothesized that, because of their role in controlling freezing, GABAergic neurons of the vlPAG are poised to present a neuronal substrate for the interaction of freezing and flight pathways. Consequently, we asked whether glutamatergic, flight-promoting neurons of the dl/IPAG could negatively regulate vlPAG excitatory output via activation of vlPAG GABAergic neurons to inhibit the freezing response. In offspring of *Vglut2-ires-Cre* crossed with *Gad1-eGFP* mice, we performed minimal injections of diluted virus to Cre-dependently express Syn-Myc in vGluT2<sup>+</sup> dl/IPAG neurons only (Fig. 5f). In line with our hypothesis, we found that dl/IPAG glutamatergic neurons form synaptic contacts with GABAergic vlPAG cells (Fig. 5g, h), whose optical activation can lead to flight responses (Fig. 3k). These results support the notion that vlPAG GABAergic neurons integrate multiple inhibitory and excitatory inputs from distinct upstream brain areas to regulate the selection of appropriate active or passive defensive behaviours.

## Discussion

Our study defines an amygdala–midbrain–medullary circuit through which freezing behaviour, an evolutionarily conserved response to threat, is generated. Central to this process is a circuit mechanism

involving the disinhibition of vlPAG glutamatergic neurons projecting to pre-motor cells located in the Mc. Disinhibition of this vlPAG→Mc pathway is generated by a disynaptic GABAergic local micro-circuit receiving inhibitory input from CEA (Extended Data Fig. 6). It is important to note that the excitatory vlPAG to Mc pathway did not mediate concomitant analgesia, a hallmark of the general defensive response to threat. Consistent with this result, our single-unit data show that some glutamatergic neurons of the vlPAG are positively and others are negatively correlated with freezing behaviour. These findings suggest that different aspects of the defensive response, such as freezing, flight and analgesia, could be mediated by distinct glutamatergic output pathways from the PAG.

Our data suggest the existence of a dedicated vlPAG output mediating freezing. However, to ensure a rapid switch between passive and active coping with fluctuating threat levels, interactions between freezing and flight circuits are required. Active defensive behaviour including flight could be driven by different hypothalamic<sup>26,27</sup> or prefrontal<sup>43</sup> inputs, directly or via disinhibition onto dl/IPAG glutamatergic neurons, which concomitantly might block freezing behaviour by activation of GABAergic neurons controlling excitatory vlPAG output to the medulla. This notion is supported by our finding of glutamatergic inputs from dl/IPAG onto GABAergic neurons of the vlPAG. While earlier models of PAG function have emphasized its columnar organization<sup>13,17,19,34,44,45</sup>, or the existence of parallel input-output pathways mediating active or passive defensive behaviours<sup>3,12,46</sup>, the model emerging from our study supports a key role for local PAG circuitry in the integration of extrinsic inputs to ensure rapid behavioural, autonomic and endocrine adaptations in the face of threat.

A growing body of evidence suggests that the interactions of distinct types of neuron within highly organized neuronal circuits are critical for any higher brain function<sup>47–49</sup>, and that circuit dysregulation contributes to psychiatric conditions, among which fear

and anxiety-related disorders are the most prevalent<sup>4,5,50</sup>. Our study shows that similar organizational principles and functional motifs exist even within evolutionarily old, mammalian 'survival circuits'<sup>3</sup> dedicated to expression of defensive behaviours. A mechanistic functional understanding of these circuits will provide new insights into possible mechanisms underlying human psychiatric conditions associated with maladaptive coping behaviours under stressful conditions.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 10 September 2015; accepted 8 April 2016.**

**Published online 1 June 2016.**

- Anderson, D. J. & Adolphs, R. A framework for studying emotions across species. *Cell* **157**, 187–200 (2014).
- Blanchard, D. C. & Blanchard, R. J. in *Handbook of Anxiety and Fear* Vol. 17 (eds Blanchard, R. J., Blanchard, D. C., Griebel, G. & Nutt, D.) Ch. 2.4, 63–79 (Academic, 2008).
- LeDoux, J. Rethinking the emotional brain. *Neuron* **73**, 653–676 (2012).
- Deisseroth, K. Circuit dynamics of adaptive and maladaptive behaviour. *Nature* **505**, 309–317 (2014).
- Tovote, P., Fadok, J. P. & Lüthi, A. Neuronal circuits for fear and anxiety. *Nature Rev. Neurosci.* **16**, 317–331 (2015).
- Rosen, J. B. & Schulkin, J. From normal fear to pathological anxiety. *Psychol. Rev.* **105**, 325–350 (1998).
- Perusini, J. N. & Fanselow, M. S. Neurobehavioral perspectives on the distinction between fear and anxiety. *Learn. Mem.* **22**, 417–425 (2015).
- Blanchard, D. C., Griebel, G., Pobbe, R. & Blanchard, R. J. Risk assessment as an evolved threat detection and analysis process. *Neurosci. Biobehav. Rev.* **35**, 991–998 (2011).
- Blanchard, R. J. & Blanchard, D. C. Crouching as an index of fear. *J. Comp. Physiol. Psychol.* **67**, 370–375 (1969).
- Fanselow, M. S. & Bolles, R. C. Naloxone and shock-elicited freezing in the rat. *J. Comp. Physiol. Psychol.* **93**, 736–744 (1979).
- Blanchard, D. C., Griebel, G. & Blanchard, R. J. Mouse defensive behaviors: pharmacological and behavioral assays for anxiety and panic. *Neurosci. Biobehav. Rev.* **25**, 205–218 (2001).
- Gross, C. T. & Canteras, N. S. The many paths to fear. *Nature Rev. Neurosci.* **13**, 651–658 (2012).
- Carrive, P. The periaqueductal gray and defensive behavior: functional representation and neuronal organization. *Behav. Brain Res.* **58**, 27–47 (1993).
- Leman, S., Dielenberg, R. A. & Carrive, P. Effect of dorsal periaqueductal gray lesion on cardiovascular and behavioural responses to contextual conditioned fear in rats. *Behav. Brain Res.* **143**, 169–176 (2003).
- Tomaz, C., Brandão, M., Bagri, A., Carrive, P. & Schmitt, P. Flight behavior induced by microinjection of GABA antagonists into periventricular structures in detelencephalated rats. *Pharmacol. Biochem. Behav.* **30**, 337–342 (1988).
- Walker, P. & Carrive, P. Role of ventrolateral periaqueductal gray neurons in the behavioral and cardiovascular responses to contextual conditioned fear and poststress recovery. *Neuroscience* **116**, 897–912 (2003).
- Zhang, S. P., Bandler, R. & Carrive, P. Flight and immobility evoked by excitatory amino acid microinjection within distinct parts of the subnucleus midbrain periaqueductal gray of the cat. *Brain Res.* **520**, 73–82 (1990).
- Morgan, M. M. & Clayton, C. C. Defensive behaviors evoked from the ventrolateral periaqueductal gray of the rat: comparison of opioid and GABA disinhibition. *Behav. Brain Res.* **164**, 61–66 (2005).
- Fanselow, M. S. in *The Midbrain Periaqueductal Gray Matter: Functional, Anatomical and Immunohistochemical Organization* (eds Depaulis, A. & Bandler, R.) 151–173 (Plenum, 1991).
- Bandler, R., Depaulis, A. & Vergnes, M. Identification of midbrain neurones mediating defensive behaviour in the rat by microinjections of excitatory amino acids. *Behav. Brain Res.* **15**, 107–119 (1985).
- Bandler, R. & Carrive, P. Integrated defence reaction elicited by excitatory amino acid microinjection in the midbrain periaqueductal grey region of the unrestrained cat. *Brain Res.* **439**, 95–106 (1988).
- Tovote, P. et al. Activation of central CRF receptor 1 by cortagine results in enhanced passive coping with a naturalistic threat in mice. *Psychoneuroendocrinology* **35**, 887–895 (2010).
- Rizvi, T. A., Ennis, M., Behbehani, M. M. & Shipley, M. T. Connections between the central nucleus of the amygdala and the midbrain periaqueductal gray: topography and reciprocity. *J. Comp. Neurol.* **303**, 121–131 (1991).
- Oka, T., Tsumori, T., Yokota, S. & Yasui, Y. Neuroanatomical and neurochemical organization of projections from the central amygdaloid nucleus to the nucleus reticulospinalis via the periaqueductal gray in the rat. *Neurosci. Res.* **62**, 286–298 (2008).
- Canteras, N. S., Simerly, R. B. & Swanson, L. W. Projections of the ventral premammillary nucleus. *J. Comp. Neurol.* **324**, 195–212 (1992).
- Silva, B. A. et al. Independent hypothalamic circuits for social and predator fear. *Nature Neurosci.* **16**, 1731–1733 (2013).
- Wang, L., Chen, I. Z. & Lin, D. Collateral pathways from the ventromedial hypothalamus mediate defensive behaviors. *Neuron* **85**, 1344–1358 (2015).
- An, X., Bandler, R., Ongür, D. & Price, J. L. Prefrontal cortical projections to longitudinal columns in the midbrain periaqueductal gray in macaque monkeys. *J. Comp. Neurol.* **401**, 455–479 (1998).
- Choi, J. S. & Kim, J. J. Amygdala regulates risk of predation in rats foraging in a dynamic fear environment. *Proc. Natl Acad. Sci. USA* **107**, 21773–21777 (2010).
- Behbehani, M. M. Functional characteristics of the midbrain periaqueductal gray. *Prog. Neurobiol.* **46**, 575–605 (1995).
- Depaulis, A., Morgan, M. M. & Liebeskind, J. C. GABAergic modulation of the analgesic effects of morphine microinjected in the ventral periaqueductal gray matter of the rat. *Brain Res.* **436**, 223–228 (1987).
- Vaughan, C. W., Ingram, S. L., Connor, M. A. & Christie, M. J. How opioids inhibit GABA-mediated neurotransmission. *Nature* **390**, 611–614 (1997).
- Ciocchi, S. et al. Encoding of conditioned fear in central amygdala inhibitory circuits. *Nature* **468**, 277–282 (2010).
- LeDoux, J. E. Emotion circuits in the brain. *Annu. Rev. Neurosci.* **23**, 155–184 (2000).
- Knobloch, H. S. et al. Evoked axonal oxytocin release in the central amygdala attenuates fear response. *Neuron* **73**, 553–566 (2012).
- Wall, N. R., Wickersham, I. R., Cetin, A., De La Parra, M. & Callaway, E. M. Monosynaptic circuit tracing *in vivo* through Cre-dependent targeting and complementation of modified rabies virus. *Proc. Natl Acad. Sci. USA* **107**, 21848–21853 (2010).
- Fenno, L. E. et al. Targeting cells with single vectors using multiple-feature Boolean logic. *Nature Methods* **11**, 763–772 (2014).
- Herry, C. et al. Switching on and off fear by distinct neuronal circuits. *Nature* **454**, 600–606 (2008).
- Lima, S. Q., Hromádka, T., Znamenskiy, P. & Zador, A. M. PINP: a new method of tagging neuronal populations for identification during *in vivo* electrophysiological recording. *PLoS ONE* **4**, e6099 (2009).
- Esposito, M. S., Capelli, P. & Arber, S. Brainstem nucleus MdV mediates skilled forelimb motor tasks. *Nature* **508**, 351–356 (2014).
- Orlovsky, G. N., Deliagina, T. G. & Grillner, S. in *Neuronal Control of Locomotion: From Mollusc to Man* Ch. 10 and 12 (Oxford Univ. Press, 1999).
- Stamatakis, A. M. et al. A unique population of ventral tegmental area neurons inhibits the lateral habenula to promote reward. *Neuron* **80**, 1039–1053 (2013).
- Halladay, L. R. & Blair, H. T. Distinct ensembles of medial prefrontal cortex neurons are activated by threatening stimuli that elicit excitation vs. inhibition of movement. *J. Neurophysiol.* **114**, 793–807 (2015).
- Keay, K. A. & Bandler, R. Parallel circuits mediating distinct emotional coping reactions to different types of stress. *Neurosci. Biobehav. Rev.* **25**, 669–678 (2001).
- Depaulis, A., Keay, K. A. & Bandler, R. Longitudinal neuronal organization of defensive reactions in the midbrain periaqueductal gray region of the rat. *Exp. Brain Res.* **90**, 307–318 (1992).
- Vianna, D. M., Landeira-Fernandez, J. & Brandão, M. L. Dorsolateral and ventral regions of the periaqueductal gray matter are involved in distinct types of fear. *Neurosci. Biobehav. Rev.* **25**, 711–719 (2001).
- Letzkus, J. J. et al. A disinhibitory microcircuit for associative fear learning in the auditory cortex. *Nature* **480**, 331–335 (2011).
- Wolff, S. B. et al. Amygdala interneuron subtypes control fear learning through disinhibition. *Nature* **509**, 453–458 (2014).
- Hangya, B., Pi, H. J., Kvitsiani, D., Ranade, S. P. & Kepecs, A. From circuit motifs to computations: mapping the behavioral repertoire of cortical interneurons. *Curr. Opin. Neurobiol.* **26**, 117–124 (2014).
- Lüthi, A. & Lüscher, C. Pathological circuit function underlying addiction and anxiety disorders. *Nature Neurosci.* **17**, 1635–1643 (2014).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank C. Müller, J. Lüdke, K. Bylund, J. Alonso, T. Lu, P. Argast and P. Buchmann for technical assistance, J. J. Letzkus for input on the manuscript, and all members of the Lüthi and Arber laboratories for discussions and other help with the project. We thank L. Gelman and S. Bourke for help with microscopy, and M. Stadler for statistical advice. We are grateful to G. Keller for providing viruses for optogenetics, Z. J. Huang for initially providing the *Gad2-ires-Cre* mouse line and L. Xiao and P. Scheiffele for the anti-rabiesG antibody. This work was supported by the Novartis Research Foundation, by the National Center of Competences in Research: 'SYNAPSY — The Synaptic Bases of Mental Diseases' (financed by the Swiss National Science Foundation) as well as by a Swiss National Science Foundation Core Grant, and a European Research Council Advanced Grant to A.L. Support for S.A. and M.S.E. was provided by a European Research Council Advanced Grant, the Swiss National Science Foundation and the Kanton Basel-Stadt. P.T. and J.P.F. were supported by NARSAD Young Investigator Grants by the Brain and Behavior Foundation. M.S.E. was also supported by a long-term post-doctoral fellowship of the Human Frontier Science Program and a Synapsis Foundation Grant. F.C. and C.H. were supported by grants from the European Research Council (ERC)



under the European Union's Seventh Framework Program (FP7/2007-2013)/ERC grant agreement number 281168 and the Fondation pour la Recherche Médicale.

**Author Contributions** P.T. conceived, designed, performed and analysed most of the experiments and wrote the manuscript. M.S.E. conceived, designed, performed and analysed neuroanatomical tracing experiments. P.B. performed and analysed *in vitro* slice recordings. F.C. and C.H. performed single-unit recordings in the PAG. S.B.E.W. established optogenetic methodology. J.P.F. performed experiments. M.M. designed and tested viruses

for optogenetics. C.R., L.F. and K.D. produced viruses for optogenetics. S.A. designed viruses for tracing and optogenetics. A.L. conceived the project and wrote the manuscript. All authors contributed to the experimental design and commented on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.T. ([philip.tovote@fmi.ch](mailto:philip.tovote@fmi.ch)) or A.L. ([andreas.luthi@fmi.ch](mailto:andreas.luthi@fmi.ch)).

## METHODS

**Animals.** Experimental subjects were adult (2- to 5-month-old) male, wild-type (Charles River Laboratories) or mutant mice of the C57BL/6J strain. *Slc17a6<sup>tm2(cre)Lowl</sup>* (*Vglut2-ires-Cre*) and *Chat<sup>tm2(cre)Lowl</sup>* (*Chat-ires-Cre*) mice were obtained from Jackson Laboratories. Founders for a *Gad2<sup>tm2(cre)Zjh</sup>* (*Gad2-ires-Cre*) and *Gad1-eGFP* mice colony were initially provided by Z. J. Huang. *Tau-lox-stop-lox-SynGFP-IRES-nlsLacZpA* mice came from an in-house colony<sup>40</sup>. All mice were individually housed in a 12 h light/dark cycle and all experiments were performed during the light cycle. Food and water was available *ad libitum*. Sample sizes were estimated based on previous studies using similar experimental designs<sup>26,33,38,40,47</sup>. All animal procedures were performed in accordance with institutional guidelines and were approved by the Veterinary Department of the Canton of Basel-Stadt.

**Viral injections and optogenetics.** Isoflurane (Attane, induction 3%, maintenance 1.5%; Provect) in oxygen-enriched air (Oxymat 3; Weinmann) was used to anaesthetize mice fixed in a stereotactic frame (Kopf Instruments 1900 series). Before opening of the scalp, local injections of ropivacain (Naropin; AstraZeneca) provided analgesia during surgery. After completion of surgery, intraperitoneal injections of meloxicam were administered to alleviate pain (60 µl of 0.5 mg ml<sup>-1</sup>, Metacam; Boehringer). A feedback-controlled heating pad (FHC) ensured maintenance of core body temperature at 36 °C. A volume of 50–200 nl virus solution (depending on respective viral titre and observed expression strength) was pressure-injected intracranially using calibrated glass pipets (5 µl microcapillary tube; Sigma-Aldrich) connected to a picospritzer III (Parker). To avoid the subcranial midline blood sinus targeting the vIPAG, holes with a diameter of 0.3 mm were drilled bilaterally into the skull at ±1.7 mm (dI/IPAG: ±1.4 mm) from midline suture, and at the level of the lambda suture. The injection capillary was then slowly lowered using a hydraulic micropositioner (Kopf Instruments model 2650) at a zenith angle of 26° to the target depth of 3 mm (dI/IPAG: ±2.6 mm) below brain surface. Coordinates for CEA injections were –1.1 mm caudal and ±2.7 mm mediolateral to bregma, at –4.2 mm perpendicularly below brain surface. The Mc of the medulla was targeted with bilateral perpendicular injections –6.4 mm caudal and ±0.95 mm mediolateral to bregma, with an injection depth of 5.6 mm. Cell-type-specific expression of optical actuators was achieved using the following, Cre-dependent AAV: rAAV(2/5)/EF1a-flex-hChR2(T159C)-mCherry (UNC Vector Core), rAAV(2/9)/CAG-flex-ReaChR-Citrine-YFP-WPRE (custom design, Vector Biolabs), rAAV(2/7) EF1a-flex-ChR2(H134R)-2A-NpHR-2A-Venus<sup>47</sup> and rAAV(2/5)CBA-flex-ARCH-GFP<sup>48</sup>. For characterization of the functional connection between GABAergic and glutamatergic neurons in the vIPAG, we introduced ChR2 into vGluT2<sup>+</sup>, putative GABAergic neurons in the vIPAG using a double-conditional approach in *Vglut2-ires-Cre* mice. We co-injected two AAVs: one that delivered Flp recombinase in an unconditional manner and one that mediated expression of ChR2 only in the presence of Flp and in the absence of Cre (AAVdj/hSyn-Cre<sup>OFF</sup>-Flp<sup>ON</sup>-hChR2(H134R)-eYFP)<sup>37</sup>. Visual targeting of vGluT2<sup>+</sup> neurons was achieved by Cre-dependent expression of tdTomato (rAAV(2/9)-flex-tdTomato). For Cre- and Flp- dependent expression of ChR2 in the glutamatergic vIPAG-to-Mc projection, we combined an injection of a retrogradely trafficked HSV (HSV/hEF1a-LSIL-mCherry-IRES-flpo; R. Neve) into the Mc with another injection of AAV(dj)/hSyn-Cre<sup>ON</sup>/Flp<sup>ON</sup>-hChR2(H134R)-mCherry<sup>37</sup> into the vIPAG of *Vglut2-ires-Cre* mice. Control mice were injected with the following AAVs: AAV(2/5)EF1a-flex-tdTomato (provided by G. Keller), AAV(2/9)CAG-flex-eGFP-WPRE-bGH, AAV(2/9)CAG-flex-tdTomato-WPRE-bGH (both Penn Vector) and AAV(dj)/hSyn-Cre<sup>ON</sup>/Flp<sup>ON</sup>-mCherry<sup>37</sup>.

For optical manipulation or electrophysiological recordings, mice were implanted with custom-built fibre connectors (fibre: 0.48 numerical aperture, 200 µm diameter; Thorlabs) 3–4 weeks after virus injections. The tip of the fibre was lowered at an angle of 26° to 250 µm above the injection site in the PAG. Implants were fixed to the skull with skull screws (P.A. Precision Components), cyanoacrylate glue (Ultra Gel; Henkel) and dental cement (Paladur; Heraeus). All fibre connectors were tested for effective light transduction before implantation. For optical stimulation of ChR2, laser light of 473 nm (CNI Laser) was applied, whereas laser light of 594 nm was used for optical stimulation of NpHR or Arch. Light intensity was adjusted with an optical power meter (Thorlabs) to reach 10–15 mW at the end of the implanted fibre stub.

**Histology.** After completion of experiments, mice were transcardially perfused with 4% paraformaldehyde in phosphate-buffered saline (PBS). Fixed brains were cryoprotected in 30% sucrose/PBS and cut on a cryostat in 80 µm coronal slices. Antibody stainings were performed on single-well floating tissue sections. Sections were incubated for 48 h in primary antibodies at 4 °C followed by one overnight incubation with secondary antibodies at 4 °C. Primary antibodies used in this study were as follows: chicken anti-GFP 1:1,000 (A10262, Molecular Probes), rabbit anti-RFP 1:5,000 (600-401-379, Rockland Immunochemicals), guinea pig anti-vGluT2 1:5,000 (AB5907, Chemicon), mouse anti-V5 1:1,000 (R960CUS, Invitrogen),

mouse anti-NeuN 1:1,000 (MAB377, Chemicon), mouse anti-Myc 1:100 (CRL-1729, ATCC), goat anti-Bgal 1:4,000 (AR2282, Biogenesis), mouse anti-channelrhodopsin-2 1:2 (clone 15E2, 651180, PROGEN Biotechnik), guinea pig anti-rabiesG 1:500 (provided by P. Scheiffele). Fluorophore-tagged secondary antibodies used were Alexa Fluor 488 donkey anti-chicken IgY 1:1,000 (703-545-155, Jackson), Cy3 donkey anti-rabbit IgG 1:1,000 (711-165-152, Jackson), Alexa Fluor 657 donkey anti-guinea pig IgG 1:1,000 (706-605-148, Jackson), Alexa Fluor 657 donkey anti-mouse IgG 1:1,000 (715-605-150, Jackson), Alexa Fluor 657 donkey anti-goat IgG 1:1,000 (705-605-147, Jackson) or Alexa Fluor 488 donkey anti-mouse IgG 1:1,000 (A21202, Molecular Probes). For counterstaining, sections were incubated for 10 min with 4',6-diamidin-2-phenylindol (DAPI, 1:10,000, Sigma). Stained brain sections were mounted on gelatin-coated slides and coverslipped with custom-made glycerol-based medium (Fluorostab). Slides were imaged using an automated slide scanner microscope (Zeiss AxioScan).

Placement of the optical fibres was assessed on the basis of the lesion the fibre tip inside the brain tissue. Mice with no or unilateral expression of the virus, or fibre tip placement outside of the PAG, were excluded from the analysis. To analyse virus expression in *Vglut2-ires-Cre* mice expressing ChR2(T159C), we outlined the area of somatic viral expression on the respective sections in a mouse brain atlas<sup>51</sup> for each animal, and then overlaid the areas at 30% transparency (Adobe Illustrator) to visualize the average centre of expression in each behavioural group.

**Behaviour.** Mice of different litters but the same genotype were used in individual experiments. No criteria were used to allocate mice to experimental groups, and, for blinding, experimental subjects had unique letter/number identifiers that indicated genotype but no group assignment. All basic testing of light-induced behavioural effects was performed under low-fear conditions in a novel, circular Plexiglas cylinder with smooth white floor (diameter 27 cm) under dim-light conditions in a dark-walled sound-attenuated chamber. Acetic acid (1%) was used to clean the context and to provide a distinct olfactory stimulus. After 1 min of habituation, light was applied four times for 5 s, with an inter-stimulus interval of 55 s. Total duration of freezing during the four 'light on' periods was compared with time spent freezing in the period of equivalent length immediately before onset of the first light stimulus.

To investigate light-mediated effects on conditioned freezing, mice were subjected to auditory fear conditioning in a brightly illuminated square context (27 cm × 27 cm) with a metal grid floor. A train of 20 tone beeps (7.5 kHz, 75 dB sound pressure level, 500 ms duration, 500 ms inter-beep-interval) was used as the conditioned stimulus (CS) and an electrical foot-shock (0.6 mA dc, 1 s duration) was used as the unconditioned stimulus (US). The conditioning session lasted 440 s during which the mice were exposed to three back-to-back CS-US pairings in a pseudo-random fashion, with a baseline period of 180 s and a minimal inter-stimulus interval of 80 s. On the day after conditioning, mice were exposed to four CS-only presentations in a dimly illuminated context different from the conditioning context. While the conditioning context was cleaned with 70% ethanol, the retrieval context was wiped down with 1% acetic acid. The duration of the retrieval session was 540 s, with a baseline period of 180 s and a pseudo-random presentation of the CS with a minimal inter-stimulus interval of 120 s. To test for light-activation effects on CS-induced defensive behaviour, the second and fourth presentation of the CS was paired with 20 s of continuous laser light. Total duration of freezing during the two CS-alone periods (first and third) was compared with time spent freezing during the two CS periods (second and fourth) paired with 'light on', and with a baseline period of equal length (40 s) directly before onset of the first CS. Contextual fear was tested a day later by re-introducing the experimental subject into the original conditioning context for 5 min. To test for effects of light activation on contextual freezing laser illumination was turned on twice for 1 min, with a 1 min pre-baseline and a 1 min inter-stimulus interval. Total duration of freezing during the two 'light on' periods was compared with time spent freezing in the time period of equivalent length immediately before light onset.

For *in vivo* recordings of unidentified PAG single units, auditory fear conditioning and testing took place in two different contexts (contexts A and B). To measure movement, an automated infrared beam detection system located on the bottom of the experimental chambers was used (Coulbourn Instruments). The animals were considered to be freezing if no movement was detected for 2 s (ref. 38). On day 1, C57BL6/J mice were submitted to a habituation session in context A, in which they received four presentations of the CS<sup>+</sup> and the CS<sup>-</sup> (total CS duration, 30 s; consisting of 50-ms pips at 0.9 Hz repeated 27 times, 2 ms rise and fall, pip frequency, 7.5 kHz or white-noise, 80 dB sound pressure level). Discriminative fear conditioning was performed on the same day by pairing the CS<sup>+</sup> with a US (1 s foot-shock, 0.6 mA, 5 CS<sup>+</sup>-US pairings, inter-trial intervals, 20–180 s). The onset of the US coincided with the offset of the CS<sup>+</sup>. The CS<sup>-</sup> was presented after each CS<sup>+</sup>-US association but was never reinforced (five CS<sup>-</sup> presentations; inter-trial intervals, 20–180 s). The frequencies used for CS<sup>+</sup> and CS<sup>-</sup> were counterbalanced

across animals. On day 2, conditioned mice were submitted to a testing session (retrieval session) in context B during which they received four presentations of the CS<sup>-</sup> and CS<sup>+</sup>.

To test for effects on unconditioned freezing, mice were exposed to a remote-controlled toy snake in an open-field arena (50 cm × 50 cm). To control for baseline activity and laser effects, mice were pre-exposed to the open field alone for 10 min, with four laser illumination periods of 20 s (inter-stimulus interval 40 s) in the second half of this phase. After introduction of the snake, mice remained in the open field for another 5 min, during which the snake's movement was remotely controlled by the experimenter outside the chamber. To maintain high freezing levels, the snake was moved around the area, thereby covering varying distances to the mouse without establishing direct contact.

Locomotor activity was recorded by an infrared beam system (Coulbourn Instruments) or an overhead video tracking system (CinePlex Studio). Freezing was defined as immobility detected by lack of beam breaks for 2 s, as described before<sup>38</sup>. Using the video tracking system, freezing was extracted with the freezing detector plug-in (CinePlex Editor). A 2 s criterion of the thresholded motion measure, based on a contour-tracking algorithm, was used to define freezing. Motion measure was automatically computed as the normalized frame-by-frame difference of the animal's body contour in pixels. Automatically detected freezing behaviour was cross-checked on the video recording to exclude false-positive freezing bouts, for example during grooming episodes, or include false negative freezing intervals, for example owing to motion artefacts caused by cable movement in front of the camera. Timestamps for freezing episodes and stimulation events (CS, US, laser) were imported into data analysis software (Neuroexplorer 5, Nex Technologies) and averaged over the respective time interval. Behavioural activity of the animal was assessed using the motion measure of the video tracking system. Motion was averaged over 1-s bins and normalized as the percentage change in relation to a baseline period. Baseline was the time directly before stimulus onset: that is, 16 s before 'light on' in experiments with naive mice, and 10 s before CS/CS<sup>+</sup> 'light on' in cued fear-conditioning experiments. Combined contour tracking of the mouse and colour-tracking of the toy snake were used to extract *x-y* coordinates of the two subjects and calculate their distance.

**Tail immersion test.** To test for analgesic effects induced by optical stimulation of glutamatergic cells of the PAG, mice tail tips were immersed in hot water with a temperature of 50 °C. This was done eight times with an inter-trial interval of 40 s, and tail withdrawal latency was scored frame-by-frame (Windows Live Movie Maker) from the video recorded during the test session (Plexon Cineplex). On the last four trials, laser light was turned on for 5 s directly before the tail immersion. If the mouse did not withdraw its tail within 10 s of immersion, the trial was terminated. Tail withdrawal latencies of the first four non-manipulated trials were averaged as baseline and compared with average withdrawal latency during the last four light-stimulated trials.

**Slice electrophysiology.** Standard procedures were used to prepare 300 µm thick coronal slices from 12- to 14-week-old male *Vglut2-ires-Cre* or *Vglut2-ires-Cre::Gad1-eGFP* mice, which received intracranial virus injections 4 weeks before. The brain was dissected in ice-cold artificial cerebrospinal fluid, mounted on an agar block and sliced with a vibrating-blade microtome (HM 650 V, Carl Zeiss) at 4 °C. Slices were maintained for 45 min at 37 °C in an interface chamber containing artificial cerebrospinal fluid equilibrated with 95% O<sub>2</sub>/5% CO<sub>2</sub> and containing the following (in mM): 124 NaCl, 2.7 KCl, 2 CaCl<sub>2</sub>, 1.3 MgCl<sub>2</sub>, 26 NaHCO<sub>3</sub>, 0.4 NaH<sub>2</sub>PO<sub>4</sub>, 18 glucose, 4 ascorbate. Recordings were performed with artificial cerebrospinal fluid in a recording chamber at a temperature of 35 °C at a perfusion rate of 1–2 ml min<sup>-1</sup>. PAG neurons were visually identified with infrared video microscopy using an upright microscope equipped with a ×40 objective (Olympus). Patch electrodes (3–5 MΩ) were pulled from borosilicate glass tubing. For voltage clamp experiments to record eIPSCs, patch electrodes were filled with a solution containing the following (in mM): 110 CsCl, 30 K-gluconate, 1.1 EGTA, 10 HEPES, 0.1 CaCl<sub>2</sub>, 4 Mg-ATP, 0.3 Na-GTP (pH adjusted to 7.3 with CsOH, 280 mOsm) and 4 N-(2,6-dimethylphenylcarbamoylmethyl) triethylammonium bromide (QX-314; Tocris-Cookson).

Evoked IPSCs were elicited by 10 ms blue-light stimulation of either local vPAG axon terminals of non-glutamatergic neurons expressing ChR2 or ChR2<sup>+</sup> CEA axons projecting to PAG. To exclude glutamatergic inputs, CNQX (6-cyano-7-nitroquinoxaline-2,3-dione, 10 µM; AMPA receptor antagonist) and (R)-CPP ((RS)-3-(2-carboxypiperazin-4-yl)-propyl-1-phosphonic acid, 10 µM; NMDA receptor antagonist) were added to the artificial cerebrospinal fluid. To confirm the eIPSCs GABAergic nature, picrotoxin (100 µM) was added at the end of the recordings. Successful connections were scored if the amplitude of eIPSCs was higher than 10 pA, with the latency within 10 ms for at least 60% of the trials (six out of ten trials). Whole-cell patch-clamp recordings were excluded if the access resistance exceeded 13 MΩ and changed more than 20% during the recordings.

Data were recorded with a MultiClamp 700B (Molecular Devices) amplifier, filtered at 0.2 kHz and digitized at 10 kHz. Data were acquired and analysed with Clampex 10.0, Clampfit 10.0 (Molecular Devices). All chemicals for the internal and external solutions were purchased from Fluka/Sigma. Glutamatergic blockers were purchased from Tocris Bioscience.

**Anatomical tracing.** To characterize CEA inputs to vPAG, we injected retrogradely transported fluorescent latex beads (Lumafluor) into the vPAG of two *Gad1-eGFP* mice. Four days after injection, mice were killed, transcardially perfused with 4% paraformaldehyde in PBS, and brains were extracted and processed for histology as described above. On four coronal sections from each mouse, bead<sup>+</sup> cells in the CEI and CEM were counted and normalized against total cells stained by NeuN.

To identify the presynaptic partners of specific neuronal subpopulations in the PAG, we used a monosynaptically restricted pseudotyped rabies virus<sup>36</sup>. We performed a local injection in the vPAG or IPAG of AAVs conditionally delivering rabies glycoprotein (AAV(2/9)CAG-flex-rabiesG-2A-H2B-10xV5-tag, short: AAV-flex-rabiesG)<sup>40</sup> and TVA (AAV(2/9)CAG-flex-TVA-2A-H2B-eGFP, short: AAV-flex-TVA)<sup>40</sup> into either *Vglut2-ires-Cre* or *Gad2-ires-Cre* mice. Two weeks later, we injected EnvA-ΔG-mCherry-rabies<sup>36</sup> into the same location. Mice were killed 7 days thereafter, transcardially perfused with 4% paraformaldehyde in PBS, and brains were extracted and processed for histology as described above. Brain sections corresponding to the injection site were stained for GFP, red fluorescent protein (RFP) and V5, and counterstained with DAPI. Brain sections outside the injection site were stained for RFP and NeuN. To compare the relative input to the different PAG subpopulations, we quantified the number of starter cells within the PAG (mCherry, GFP and V5 triple-positive cells) in four sections around the injection site and we counted the number of mCherry<sup>+</sup> cells within the entire CEA. Images were acquired with an Olympus confocal microscope (FV1000) with a motorized stage, using a ×20 objective, a 3 × 3 tiled scan with 15% overlap and a step size of 1.5 µm along the depth of the slice. Triple-positive cells in the PAG and mCherry<sup>+</sup> cells in CEA were counted manually along every confocal plane using Imapar software (Bitplane). To quantify the hypothalamic input to PAG subpopulations, we analysed approximately half the volume of the hypothalamus by imaging every other section along the rostro-caudal axis with an automated slide scanning microscope (Zeiss Axioscan Z1) using a ×10 objective. mCherry<sup>+</sup> neurons were counted manually on the image projection. We use the following nomenclature for the hypothalamic nuclei: dorsomedial hypothalamic nucleus (DM), anterior hypothalamus, medial part (AHM), including the entire anterior hypothalamic area, the latero-anterior hypothalamic nucleus and the paraventricular hypothalamic nucleus), ventromedial hypothalamic nucleus (VMH), posterior hypothalamic nucleus (PH), peduncular part of lateral hypothalamus (PLH, including the Mc of the lateral hypothalamus), lateral hypothalamic area (LH, including the parasubthalamic nucleus), premammillary nucleus (PMD). All quantifications were performed by an experimenter blind to the subject's genotype.

To study PAG projections to brainstem neurons directly connected to spinal motor neurons (pre-motor neurons), we combined an anterograde vPAG injection of an AAV expressing presynaptic fluorescent markers and monosynaptic rabies spreading from spinal motor neurons. To label pre-motor neurons, we injected a monosynaptic rabies virus into forelimb muscles that retrogradely infected the corresponding spinal motor neurons. To allow for spreading of the rabies virus in adult mice, we complemented cervical motor neurons with an AAV conditionally expressing rabies glycoprotein through an intra-spinal injection in *Chat-Cre* mice. First, we injected into the vPAG of *Chat-Cre* adult mice a custom-made AAV that unconditionally expressed GFP-tagged synaptophysin (Syn) in presynaptic terminals (AAV(2/9)CAG-flex-SynGFP + AAV(2/9)CMV-Cre)<sup>40</sup>. In the same surgery session, we injected AAV(2/9)CAG-flex-rabiesG-2A-H2B-10xV5-tag<sup>40</sup> into the cervical part of the spinal cord. Two weeks thereafter, G-deleted rabies virus coated with the CVS-glycoprotein was injected in triceps and biceps muscles. Eight days after rabies injection, mice were killed and brains were immunostained against GFP, RFP and vGluT2 and counterstained with DAPI. High-resolution three-dimensional images of eight complete pre-motor cells in the Mc were acquired on a custom-made dual motorized spinning-disk microscope (Life Imaging Services) using a ×63 objective, 8 × 8 tile scan and 0.2 µm step size. We quantified the number of GFP<sup>+</sup> (vGluT2<sup>+</sup>) vPAG inputs to soma and dendritic tree of these cells manually in eight neurons from three mice. The Mc was defined as described before<sup>40</sup> and included lateral paragigantocellular, as well as ventral and alpha parts of the gigantocellular, reticular nucleus.

To assess local vPAG inputs specifically onto Mc-projecting glutamatergic neurons, we used an intersectional viral approach (Extended Data Fig. 4f–h). We injected a retrogradely transported HSV that Cre-dependently delivered rabies G-protein (HSV/HEF1α-LS1L-rabiesG; R. Neve) into the Mc of offspring from *Vglut2-ires-Cre* crossed with *Gad1-eGFP* mice. In the same surgery session, we



injected conditional AAV into the vIPAG to virally express the TVA receptor in glutamatergic cells (rAAV(2/9)CAG-flex-TVA-2A-H2B-eGFP, custom designed, Vector Biolabs). Two weeks later, we injected EnvA- $\Delta$ G-mCherry-rabies into the Mc. Mice were killed 7 days later and brains processed as described above. We then used standard immunohistochemistry to stain for GFP, RFP and rabies G-protein. High-resolution confocal images ( $\times 40$ ) of vIPAG were taken from three coronal sections of the PAG (rostral, medial, caudal) and individual cells were manually counted using Imaris software (Bitplane). Triple-positive cells were identified as starter cells, whereas GFP and RFP double-positive cells represented local GAD1<sup>+</sup> cells connected to glutamatergic Mc-projecting neurons.

We performed several controls to demonstrate the specificity of the monosynaptic rabies tracing technology. To check for the specificity of the AAV viruses delivering the TVA receptor (AAV-flex-TVA) and rabiesG protein (AAV-flex-rabiesG), we injected those viruses in the vIPAG of offspring from *Vglut2-Cre* and *Tau-lox-stop-lox-SynGFP-IRES-nlsLacZpA* reporter mouse lines. After 2 weeks, animals were killed and the injection site was cut, immunostained and co-localization of GFP or V5 with b-Gal was analysed. In an additional experiment, *Vglut2-ires-Cre* animals were injected with AAV-flex-TVA followed by EnvA- $\Delta$ G-mCherry-rabies 2 weeks later. Furthermore, to test for the leakiness of the EnvA- $\Delta$ G-mCherry-rabies, we injected this virus in the vIPAG of wild-type animals combined with latex beads. Animals were killed 1 week later and the entire PAG, amygdala and hypothalamus were cut and stained for mCherry and NeuN. Lastly, wild-type animals were injected with AAV-flex-TVA, AAV-flex-rabiesG and EnvA- $\Delta$ G-mCherry-rabies following the same protocol used for monosynaptic tracing experiments.

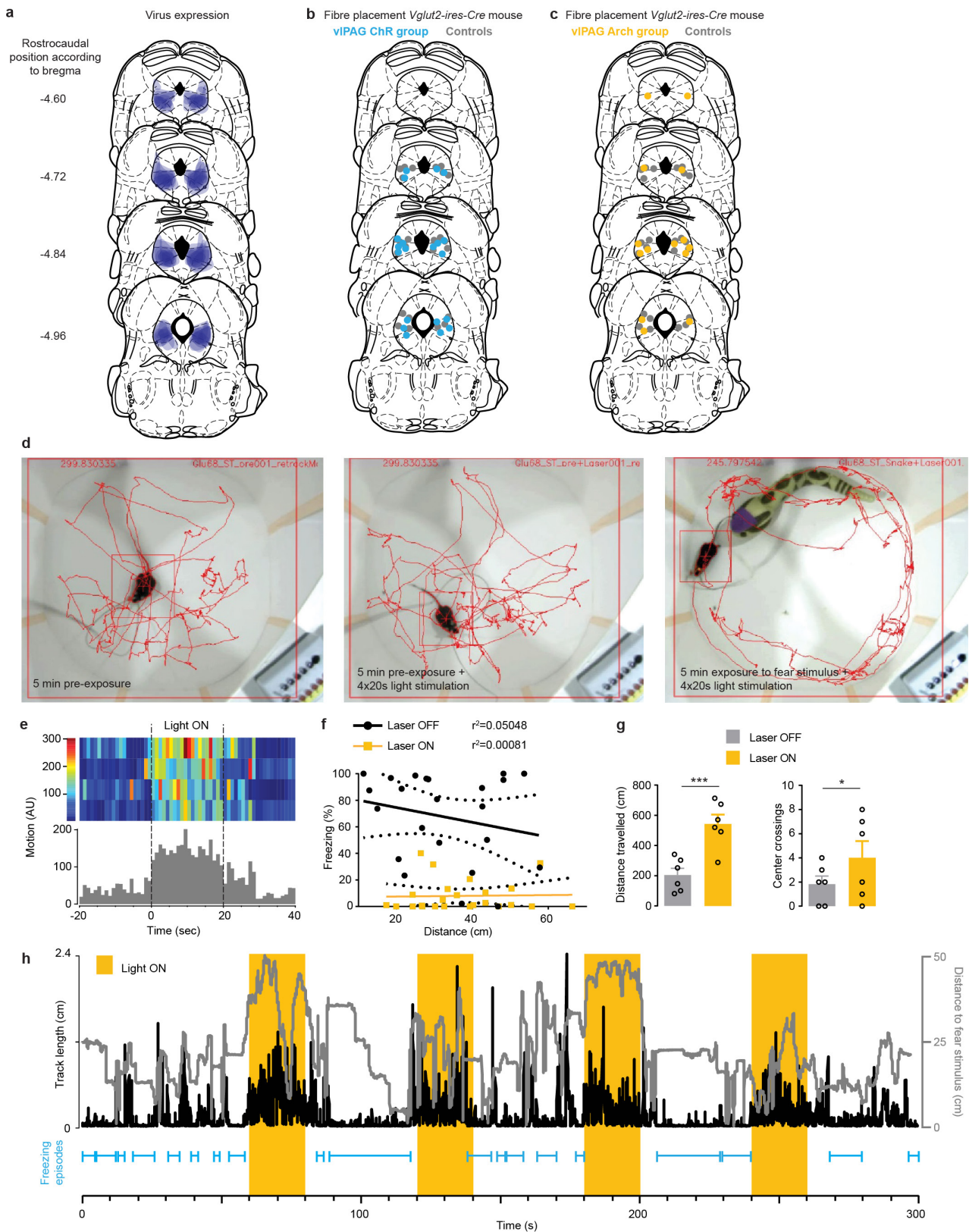
To characterize the connectivity from dIPAG to vIPAG, we injected AAV(2/9)CAG-flex-Synaptophysin-Myc<sup>40</sup> into offspring of *Vglut2-ires-Cre* mice crossed with *Gad1-eGFP* mice ( $n = 4$  mice) to visualize glutamatergic dIPAG contacts onto GFP<sup>+</sup> GABAergic vIPAG neurons. To quantify opposing synaptic contacts, we selected neurons with complete soma and acquired  $\times 60$  confocal images at a step size of 0.2  $\mu$ m with a confocal microscope (Olympus FV1000). Quantification of synaptic inputs from ipsilateral ( $n = 21$  cells from four mice) and contralateral ( $n = 13$  cells from two mice) hemispheres was performed manually using Imaris software (Bitplane).

**Single-unit recordings.** Custom-built, chronically implanted 16-wire electrodes<sup>38</sup> were used to record electrical activity in the vIPAG. Electrodes were connected to a headstage (Plexon) containing 16 unity-gain operational amplifiers. The headstage was connected to a 16-channel preamplifier (gain  $100 \times$  bandpass filter from 150 Hz to 9 kHz for unit activity, Plexon). Spiking activity was digitized at 40 kHz, bandpass filtered from 250 Hz to 8 kHz, and isolated by time-amplitude window

discrimination and template matching using a Neural Data Acquisition System (Omniplex, Plexon). Single-unit spike sorting was performed using Off-Line Spike Sorter (OFSS, Plexon) for all behavioural sessions. Principal-component scores were calculated for unsorted waveforms and plotted in a three-dimensional principal-component space; clusters containing similar valid waveforms were manually defined. A group of waveforms were considered to be generated from a single neuron if the waveforms formed a discrete, isolated, cluster in the principal-component space and did not contain a refractory period less than 1 ms, as assessed using auto-correlogram analyses. To avoid analysis of the same neuron recorded on different channels, we computed cross-correlation histograms. If a target neuron presented a peak of activity at a time that the reference neuron fired, only one of the two neurons was considered for further analysis. Spike timestamps were analysed (Neuroexplorer 5, Nex Technologies) to calculate average firing rates and z-score transformations of cells depending on behavioural parameters: that is, within and outside freezing episodes. Optical identification of single units was performed as previously described<sup>48</sup>. Briefly, laser light pulses of 100–300 ms duration were used to evoke spiking activity. Short latencies (GAD2<sup>+</sup> cells:  $\leq 15$  ms; vGluT2<sup>+</sup> cells:  $\leq 10$  ms) of reliable light-evoked spiking were considered to indicate direct light activation and, thus, allowed for identification of the cell type. To correlate single-unit activity with freezing behaviour, we calculated z-scores 2 s before and after onset of freezing during the entire recall session for recordings both of unidentified or of optically identified single units.

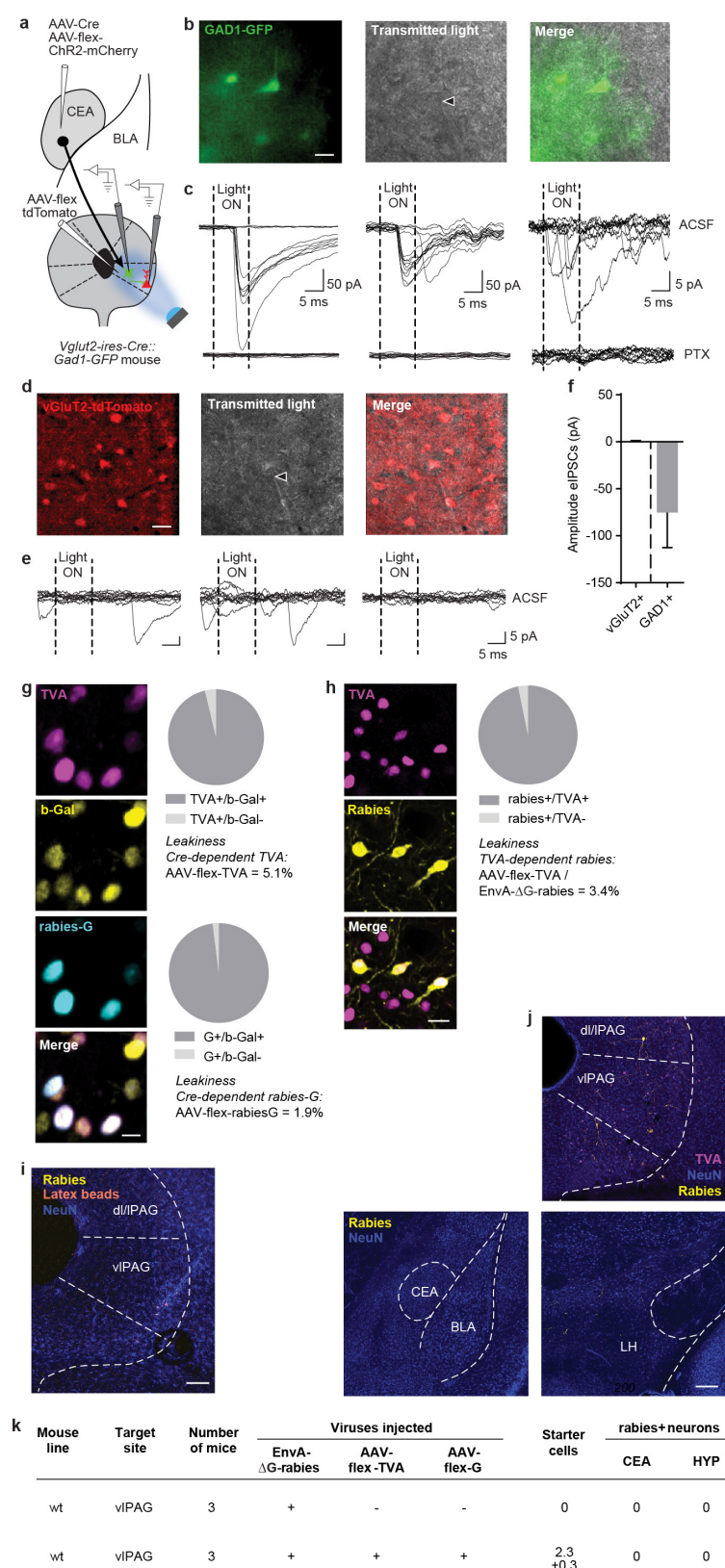
**Statistics.** The experiments were not randomized. No statistical methods were used to predetermine sample size. Data presented in box-whisker plots indicate medians, interquartile range and 5th–95th percentiles. Motion data are presented as s.e.m. range. All other data are presented as means  $\pm$  s.e.m. Statistical analyses were performed in Graphpad Prism 6.0a or using R. Normality was assessed using Shapiro-Wilk tests. Whenever the normality test failed, non-parametric Mann-Whitney or Wilcoxon signed-rank (for repeated measures) tests were used for pairwise comparisons. Within-subject group analysis of non-parametric data was performed using Friedman's test with a post-hoc Dunn's multiple comparisons test. Between-subject group analysis of non-parametric data was done with Kruskal-Wallis statistics and a post-hoc Dunn's multiple comparisons test. Variance in normally distributed data sets was analysed with one-way ANOVA and Tukey's or Sidak's post-hoc tests. Significance levels are indicated as follows: \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ . See Supplementary Information for statistics table.

51. Franklin, K. B. J. & Paxinos, G. *Atlas of the Mouse Brain* 4th edn (Academic, 2001).



**Extended Data Figure 1** | **a**, Expression of ChR2 throughout vPAG in consecutive coronal brain sections<sup>51</sup>. **b**, **c**, Fibre placements in *Vglut2-ires-Cre* mice of experimental and control groups. **d**, Supplementary Video stills with superimposed representative movements tracks during the snake open-field test for unconditioned defensive responses. **e**, Colour-coded plot for a mouse's motion before, during and after light-mediated inhibition of vPAG glutamatergic neurons expressing Arch during a snake open-field test session. **f**, Freezing responses plotted against mouse-snake distance during 'light on' and 'light off' periods showing no linear

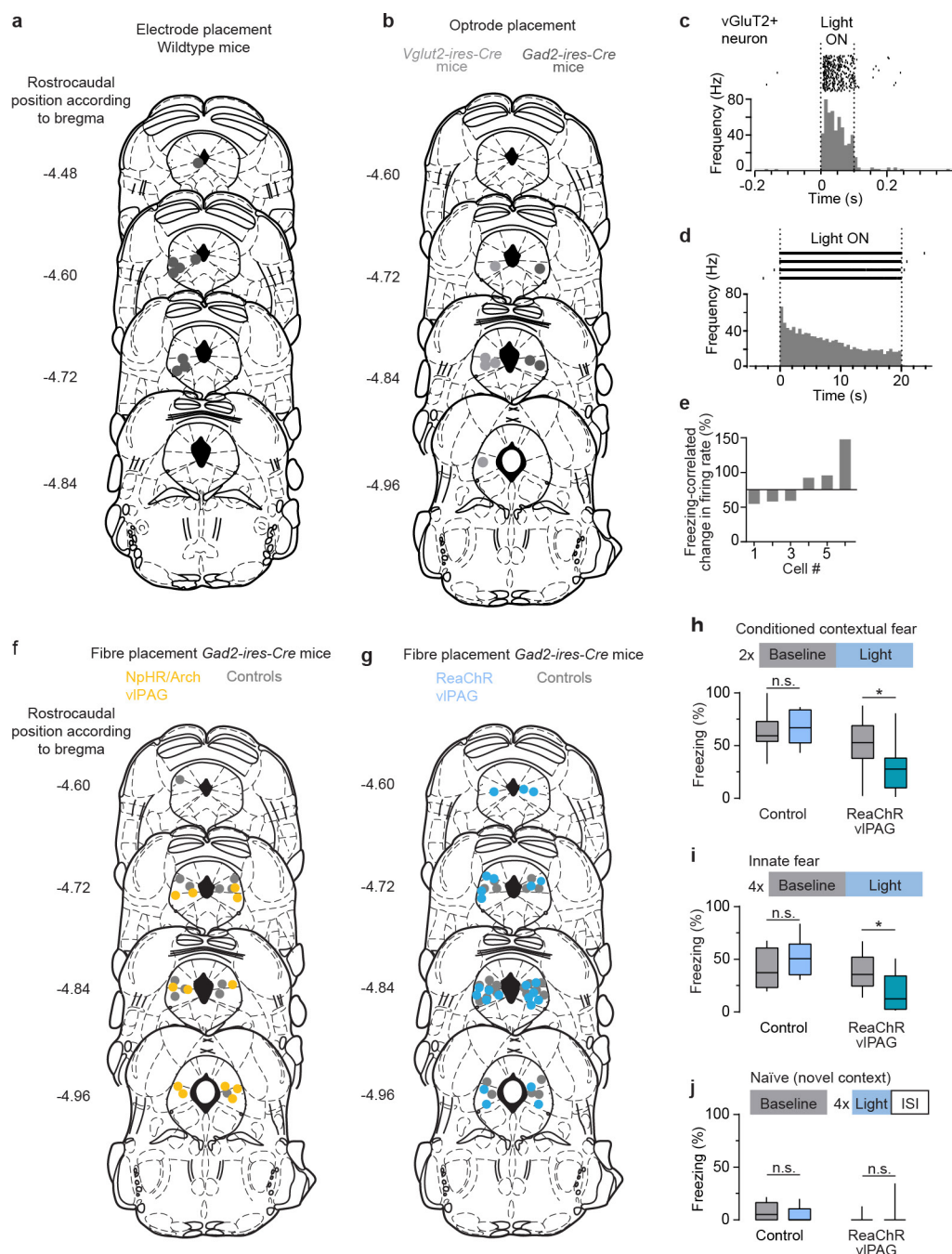
correlation. **g**, Effects of light-mediated inhibition of vPAG glutamatergic neurons on anxiety-like behaviour in the open field test with no snake present. Inhibition of vPAG glutamatergic cells resulted in enhanced track length ( $n = 6$  mice, paired two-tailed Student's  $t$ -test) and more frequent visits to the centre of the open field ( $n = 6$  mice, paired two-tailed Student's  $t$ -test). **h**, Example of an entire snake open-field test session, with track length, freezing episodes and mouse-snake distance. Values are means  $\pm$  s.e.m.



**Extended Data Figure 2 | a**, *In vitro* slice recordings of functional connectivity between CEA and GABAergic or glutamatergic neurons in vIPAG. **b**, Targeting of fluorescently labelled GAD1<sup>+</sup> neurons for whole-cell patch clamp recordings (triangle indicates patch pipette; scale bar, 20 μm). **c**, Example traces of three GAD1<sup>+</sup> cells (from three slices of three mice) showing eIPSCs upon light stimulation (10 ms duration) of afferents from the CEA (upper traces), and blockage of eIPSCs by PTX (lower traces). **d**, Targeting of fluorescently labelled vGluT2<sup>+</sup> neurons for whole-cell patch clamp recordings (triangle indicates patch pipette; scale

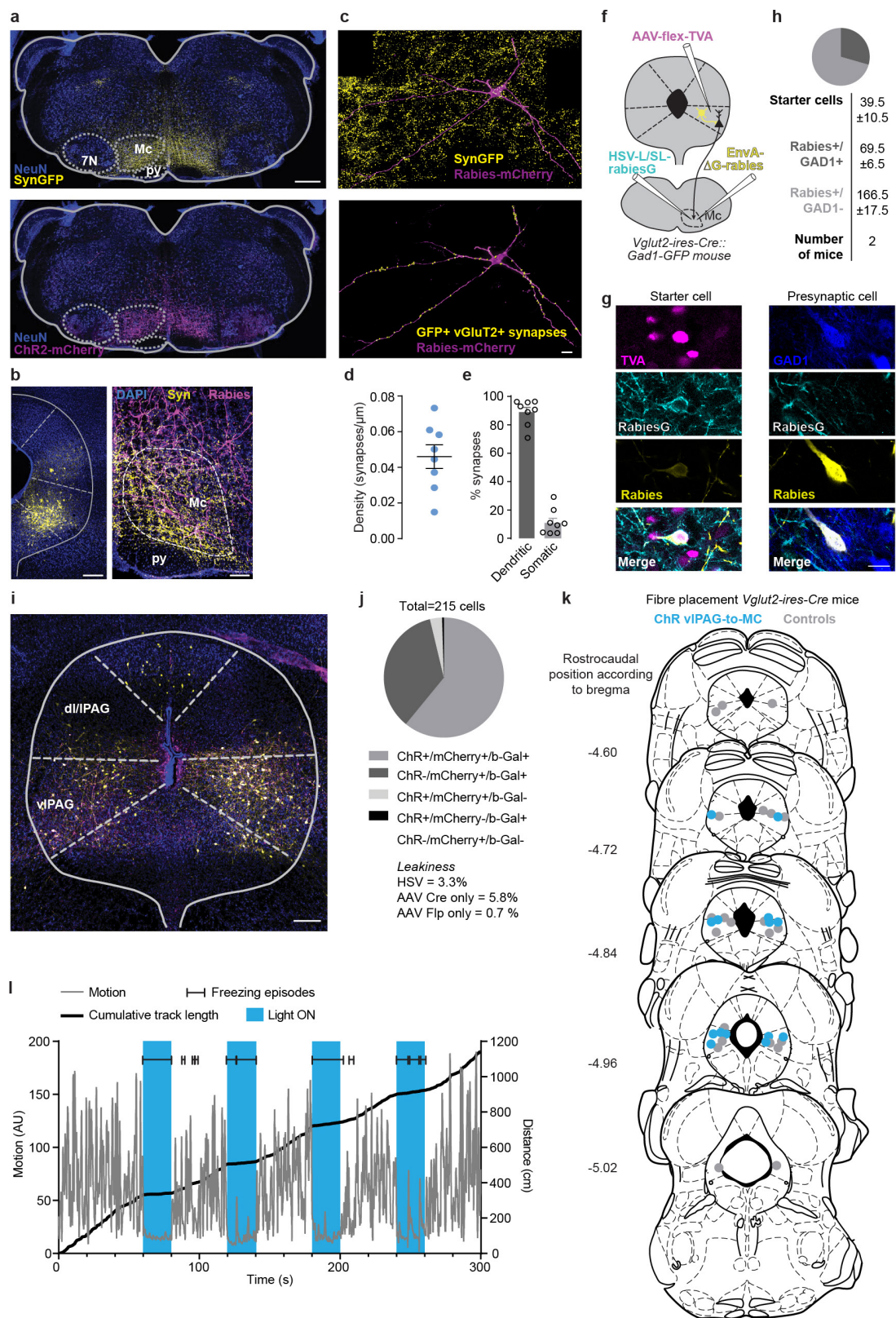
bar, 20 μm). **e**, Example traces of three light non-responsive vGluT2<sup>+</sup> neurons (from the same three slices as in **c**). **f**, Quantification of eIPSCs amplitudes. **g**, Analysis of specificity of AAV-mediated expression of TVA and rabiesG (scale bar, 10 μm). **h**, Analysis of specificity of TVA-dependent EnvA-ΔG-mCherry-rabies infection (scale bar, 20 μm). **i**, **k**, Lack of EnvA-ΔG-mCherry-rabies infection after injection into wild-type mice (scale bar, 100 μm). **j**, **k**, Leakiness analysis of the combined AAV and EnvA-ΔG-mCherry-rabies tracing system in wild-type mice (scale bar, 200 μm). Values are means ± s.e.m.





**Extended Data Figure 3** | **a**, Electrode placements for the recordings of unidentified single units. **b**, Placements of optrodes for the recordings of identified neurons. **c**, Raster-frequency plot of an optically identified vIPAG vGluT2<sup>+</sup> neuron. **d**, A glutamatergic neuron exhibiting marked optical activation during constant illumination for 20 s. **e**, Identified vGluT2<sup>+</sup> neurons ( $n = 6$ ) showed both increased and decreased activity during freezing. **f**, **g**, Fibre placements in *Gad2-ires-Cre* mice expressing inhibitory or excitatory optical actuators. **h**, **i**, Activation of vIPAG

GAD2<sup>+</sup> neurons resulted in reduced conditioned contextual freezing ( $n = 12$  ChR2,  $n = 7$  control, paired two-tailed Student's  $t$ -test) and lower innate freezing levels during the snake open-field test ( $n = 10$  ChR2,  $n = 8$  control, two-tailed Wilcoxon signed-rank test). **j**, Optical activation of vIPAG GAD2<sup>+</sup> neurons had no effect on freezing in naive mice ( $n = 8$  ChR2,  $n = 8$  control, two-tailed Wilcoxon signed-rank test). Box-whisker plots indicate median, interquartile range, and 5th–95th percentiles of the distribution. \* $P < 0.05$ .

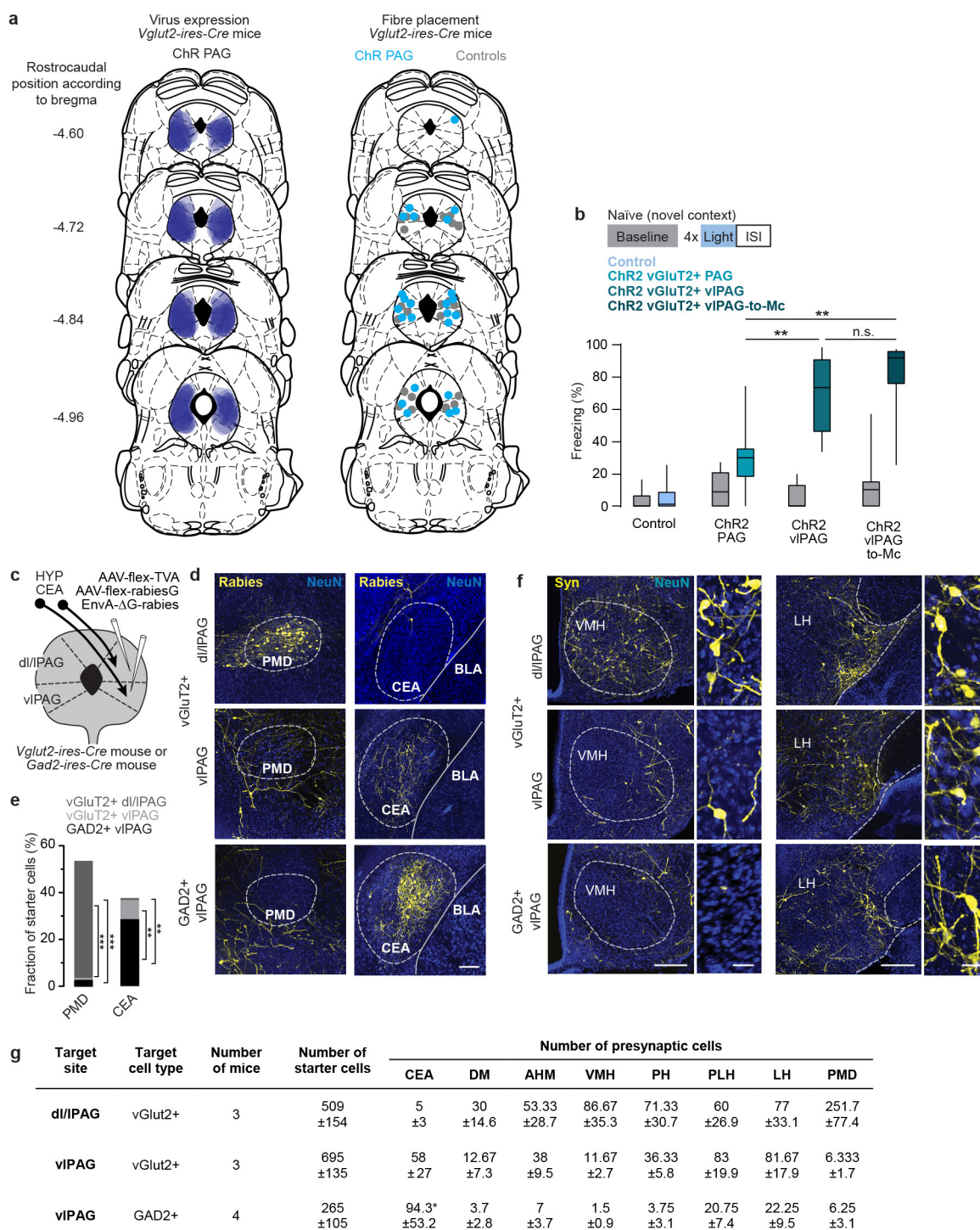


Extended Data Figure 4 | See next page for caption.

**Extended Data Figure 4 | a**, Projection pattern of glutamatergic vIPAG axonal inputs to the rostral medulla. Terminals of vGluT2<sup>+</sup> vIPAG projection neurons were labelled by AAV-mediated expression of GFP fused to presynaptic marker synaptophysin (top; scale bar, 400  $\mu$ m), and ChR2-mCherry expression was visualized using immunohistochemistry (bottom). **b**, Concomitant AAV-mediated expression of Syn-GFP in vIPAG neurons (left panel, scale bar, 200  $\mu$ m) labelled presynaptic terminals within Mc (right panel, scale bar, 100  $\mu$ m). **c**, High-resolution image of a retrogradely traced Mc pre-motor neuron (rabies-mCherry) and SynGFP<sup>+</sup> vIPAG inputs (top), and visualization of identified glutamatergic synaptic contacts (bottom; scale bar, 10  $\mu$ m). **d, e**, Density of vGluT2<sup>+</sup> vIPAG synaptic inputs to pre-motor neurons in Mc ( $n = 8$  cells from three mice), and quantification of their distribution between the dendritic or somatic compartment. **f**, Intersectional EnvA- $\Delta$ G-mCherry-rabies tracing approach to identify local GABAergic inputs to vIPAG-to-Mc-projecting glutamatergic cells. **g**, TVA-, rabiesG- and EnvA- $\Delta$ G-

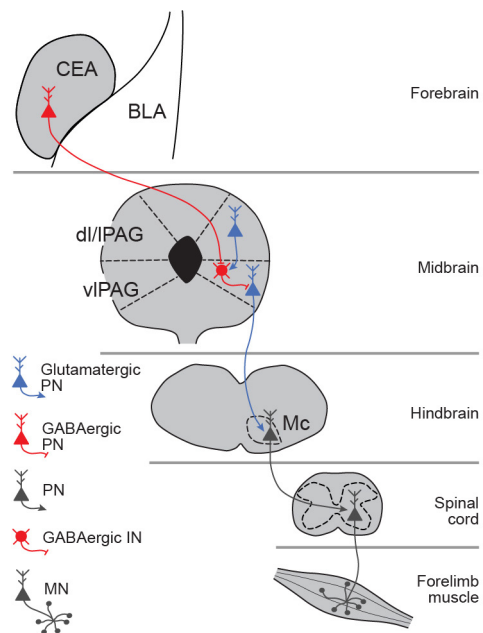
mCherry-rabies triple-positive cells were identified as starter cells (left panels), while GAD1 and EnvA- $\Delta$ G-mCherry-rabies double positive cells indicated presynaptic GABAergic neurons (right panels; scale bar, 20  $\mu$ m). **h**, Quantification of GAD1<sup>+</sup> or GAD1<sup>-</sup> presynaptic cells ( $n = 2$  mice; dark grey, rabies<sup>+</sup>/GAD1<sup>+</sup>; light grey, rabies<sup>+</sup>/GAD1<sup>-</sup>). **i**, Example picture of glutamatergic vIPAG neurons retrogradely traced from the Mc, expressing ChR2 in presence of Cre and Flp recombinase (scale bar, 200  $\mu$ m). **j**, Analysis of viral efficacy and leakiness. Overlaps of ChR delivered by AAV-Cre<sup>ON</sup>Flp<sup>ON</sup>-ChR2, HSV-delivered Cre-dependent Flp-mCherry and Cre-dependent  $\beta$ -Gal were quantified in *Vglut2::LacZ* reporter mice. **k**, Fibre placement in *Vglut2-ires-Cre* mice expressing ChR2 in glutamatergic vIPAG-to-Mc projection neurons. **l**, Example of an entire session of light activation of glutamatergic vIPAG-to-Mc projection neurons, with the mouse's motion, cumulative track length and light-induced freezing bouts. Values are means  $\pm$  s.e.m.





**Extended Data Figure 5 | a**, Expression of ChR2 throughout PAG in consecutive coronal brain sections (left), and fibre placements in *Vglut2-ires-Cre* mice of experimental and control groups (right). **b**, Light-evoked effect on freezing behaviour induced by activation of different glutamatergic subpopulations of PAG neurons in naïve animals ( $n = 12$  PAG,  $n = 10$  vIPAG,  $n = 7$  vIPAG-to-Mc, Kruskal–Wallis test,  $P < 0.001$ , Dunn's multiple comparison post-hoc test). **c**, EnvA-ΔG-mCherry-rabies-mediated, Cre-dependent monosynaptic retrograde tracing of inputs to GAD2<sup>+</sup> and vGluT2<sup>+</sup> neurons in the vIPAG and dl/IPAG. **d**, Rabies-mediated labelling of presynaptic neurons within PMD (left panels) and CEA (right panels) of *Gad2-ires-Cre* and *Vglut2-ires-Cre* mice (scale bar, 100 μm). **e**, Statistical analysis reveals differential input to vGluT2<sup>+</sup> ( $n = 3$

mice for each vIPAG and dl/IPAG) and GAD2<sup>+</sup> ( $n = 4$  mice) neurons within vIPAG or dl/IPAG. While CEA preferentially targets GAD2<sup>+</sup> neurons of the vIPAG ( $1 \times 3$  ANOVA,  $F_{(2,6)} = 21.67$ ,  $P < 0.01$ , Tukey's post-hoc test), vGluT2<sup>+</sup> neurons of the dl/IPAG receive stronger inputs from PMD ( $1 \times 3$  ANOVA,  $F_{(2,7)} = 287$ ,  $P < 0.0001$ , Tukey's post-hoc test). **f**, Cell-type-specific monosynaptic rabies tracing of VMH and LH inputs to vGluT2<sup>+</sup> neurons in the dl/IPAG, vIPAG and vIPAG GAD2<sup>+</sup> neurons (scale bars in overview, 200 μm; in zoom-in, 25 μm). **g**, Quantification of starter cells in the PAG and presynaptic cells in CEA and hypothalamic subregions. Boxes indicate median and 25th–75th percentiles of the distribution. \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ .



**Extended Data Figure 6** | Schematic representation of the freezing pathway. PN, projection neuron; IN, interneuron; MN, motor neuron.

# Acetate mediates a microbiome–brain– $\beta$ –cell axis to promote metabolic syndrome

Rachel J. Perry<sup>1</sup>, Liang Peng<sup>1</sup>, Natasha A. Barry<sup>2,3</sup>, Gary W. Cline<sup>1</sup>, Dongyan Zhang<sup>4</sup>, Rebecca L. Cardone<sup>1</sup>, Kitt Falk Petersen<sup>1,5</sup>, Richard G. Kibbey<sup>1,6</sup>, Andrew L. Goodman<sup>2,3</sup> & Gerald I. Shulman<sup>1,4,5,6</sup>

**Obesity, insulin resistance and the metabolic syndrome are associated with changes to the gut microbiota; however, the mechanism by which modifications to the gut microbiota might lead to these conditions is unknown. Here we show that increased production of acetate by an altered gut microbiota in rodents leads to activation of the parasympathetic nervous system, which, in turn, promotes increased glucose-stimulated insulin secretion, increased ghrelin secretion, hyperphagia, obesity and related sequelae. Together, these findings identify increased acetate production resulting from a nutrient–gut microbiota interaction and subsequent parasympathetic activation as possible therapeutic targets for obesity.**

Previous studies have shown that both increases<sup>1–5</sup> and decreases<sup>6,7</sup> in plasma and faecal short-chain fatty acid (SCFA) concentrations can be associated with overfeeding, obesity and the metabolic syndrome. However, whether and how alterations in SCFAs play a causal role in the development of obesity is unknown. Because plasma SCFA concentrations may not fully represent the SCFA load presented to the body, we developed a method to measure whole-body turnover rates of acetate, propionate, and butyrate by gas chromatography–mass spectrometry (GC–MS; as described in the Supplementary Methods) and found that, in contrast to propionate and butyrate, whole-body acetate turnover as well as plasma and faecal acetate concentrations were markedly increased in insulin-resistant rats after 3 days or 4 weeks on a high-fat diet (HFD) (Fig. 1a, b and Extended Data Fig. 2a–j).

Next we sought to determine the source of the increased acetate turnover in HFD-fed rats. We measured tissue acetate concentrations and dilution of the <sup>13</sup>C-acetate label during an infusion of [<sup>13</sup>C]acetate, and found each to be increased in the luminal contents of the caecum and ascending colon, with HFD-fed rats exhibiting more than a twofold increase in total acetate in the caecum and colon as well as in the brain and in [<sup>13</sup>C] bicarbonate incorporation into [<sup>13</sup>C]acetate compared to chow-fed rats (Figs 1c, d, 2a and Extended Data Fig. 2k). In order to determine conclusively the source of the increased acetate production in HFD-fed rats, we conducted four independent *in vivo* experiments to distinguish colon lumen acetate production from production by the rest of the body: we 1) washed out the contents of the caecum and colon with a saline flush; 2) ligated the portal vein below the splenic juncture; 3) treated rats with poorly absorbed broad-spectrum oral antibiotics; and 4) performed an acute colectomy. Each of these interventions reduced whole-body acetate turnover by 75–90% (Fig. 2b–f). Together, these data strongly suggest that the gut microbiota is the source of most of the increase in endogenous acetate production in HFD-fed rats. We next showed that faecal material can generate acetate *in vitro* from [<sup>13</sup>C] glucose or [<sup>13</sup>C] fatty acids, and that boiling or irradiating the faeces prevents the production of acetate, suggesting a role for faecal microbes in generating acetate (Extended Data Fig. 2l–n). Consistent with this hypothesis, treatment of the faecal material with either or both of the broad-spectrum antibiotics vancomycin and gentamycin markedly reduced acetate production (Extended Data Fig. 2o).

## Acetate drives insulin secretion

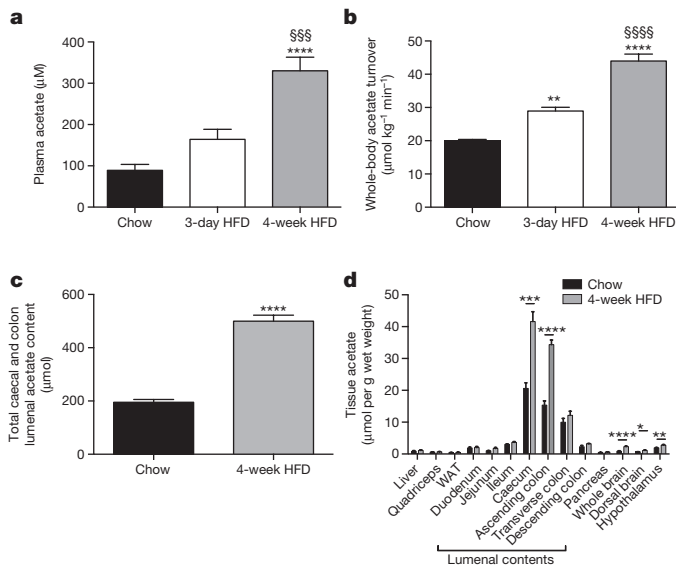
Next we examined glucose-stimulated insulin secretion (GSIS) during a hyperglycaemic clamp and measured marked increases in GSIS in 3-day and 4-week HFD-fed rats (Fig. 3a and Extended Data Fig. 3a–c). To determine whether the associated increases in acetate turnover drove this increased GSIS, we performed hyperglycaemic clamps in chow-fed rats given intra-arterial infusions of acetate to match whole-body acetate turnover to that measured in HFD-fed rats. Acetate infusion in chow-fed rats replicated the increases in GSIS measured in HFD-fed rats (Fig. 3b, c and Extended Data Fig. 3d–g), strongly implicating increased acetate turnover in driving acute increases in GSIS in HFD-fed rodents. In contrast, supplementing butyrate in chow-fed rats to match the turnover rates observed in HFD-fed rats had no effect on GSIS (Extended Data Fig. 3h–m).

To evaluate further the effects of alterations in food intake on gut acetate production, we starved 4-week HFD-fed rats for 48 h and found that this intervention resulted in ~50% reductions in whole-body acetate turnover and in GSIS; however, replacing acetate by arterial infusion of 20  $\mu\text{mol kg}^{-1} \text{min}^{-1}$  acetate resulted in restoration of GSIS in rats after 48 h food deprivation (Extended Data Fig. 4a–f). Next we performed a series of dietary interventions to assess whether simple caloric excess or variations in nutrient composition<sup>8–10</sup> were responsible for the increased acetate turnover measured in HFD-fed rats. Pair-feeding with isocaloric portions of chow or HFD produced no change in acetate turnover or GSIS, whereas dietary interventions resulting in increased caloric intake increased acetate turnover and GSIS proportionally to the total calories consumed (Extended Data Fig. 4g–n;  $R^2 = 0.90$ ). To examine the role of the gut microbiota in acetate-induced hyperinsulinaemia, we treated HFD-fed rats with broad-spectrum, poorly absorbable oral antibiotics and measured a 70% reduction in GSIS during a hyperglycaemic clamp. This reduction in GSIS was acutely reversed by infusion of acetate to match plasma acetate turnover in HFD-fed rats (Fig. 3d, e and Extended Data Fig. 4o–s).

To establish a causal relationship between the microbiota and GSIS, we next transferred faecal material from chow- or HFD-fed donor rats to chow- or HFD-fed recipients. Consistent with previous reports<sup>4,11–14</sup>, culture-independent 16S rRNA sequencing of donor

<sup>1</sup>Department of Internal Medicine, Yale University School of Medicine, New Haven, Connecticut 06520, USA. <sup>2</sup>Department of Microbial Pathogenesis, Yale University School of Medicine, New Haven, Connecticut 06510, USA. <sup>3</sup>Microbial Sciences Institute, Yale University School of Medicine, New Haven, Connecticut 06516, USA. <sup>4</sup>Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, Connecticut 06519, USA. <sup>5</sup>Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Copenhagen 2200, Denmark. <sup>6</sup>Department of Cellular & Molecular Physiology, Yale University School of Medicine, New Haven, Connecticut 06510, USA.





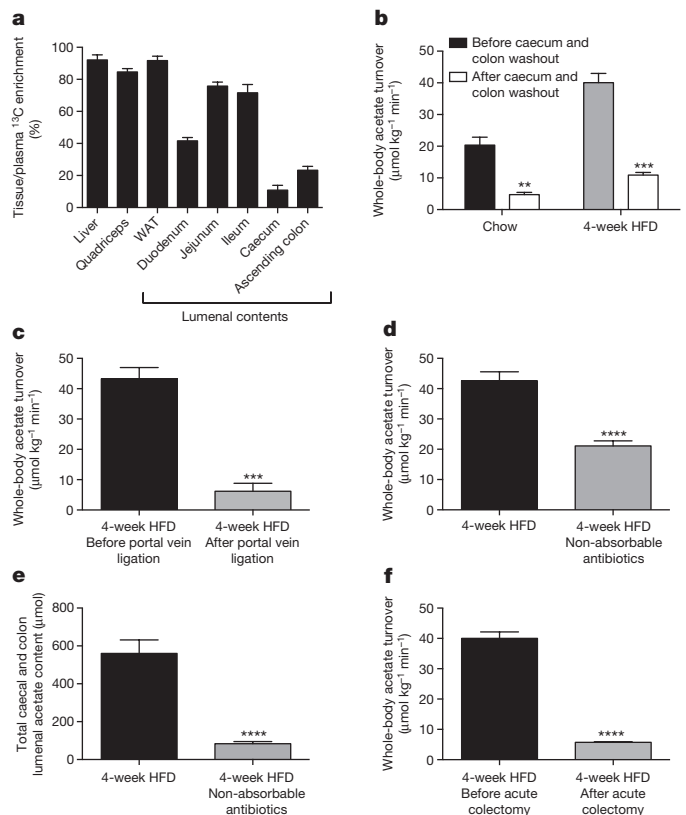
**Figure 1 | HFD-fed rats exhibit increased whole-body acetate turnover.** **a, b**, Plasma acetate concentrations and whole-body acetate turnover in chow-fed and 3-day or 4-week HFD-fed rats. **c**, Acetate content in the entire caecum and colon lumen. **d**, Tissue acetate concentrations. In all panels,  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ ,  $****P < 0.0001$  versus chow-fed rats;  $$$$P < 0.001$ ,  $$$$$P < 0.0001$  versus 3-day HFD-fed rats by one-way ANOVA with Bonferroni's multiple comparisons test (**a, b**) or by two-tailed unpaired Student's *t*-test (**c, d**). Data are the mean  $\pm$  s.e.m. of  $n = 6$  animals per group.

and recipient faecal microbiomes revealed an increase in the relative abundance of bacteria belonging to the phylum Firmicutes and a decrease in the relative abundance of representatives of the phylum Bacteroidetes in fresh faecal pellets from HFD-fed donors relative to chow-fed donors, and faecal transplantation altered the recipient animal microbiome to resemble that of the donor (Extended Data Fig. 5a–f). Notably, these faecal transplantations also transferred the corresponding acetate turnover, faecal acetate, and GSIS from the donor group to the recipient group (Fig. 3f, g and Extended Data Fig. 6a–e). However, transplantation of microbiota from chow-fed donors into chow-fed recipients by an identical procedure did not alter microbiota or metabolic phenotypes (Fig. 3f, g and Extended Data Figs 5a–f, 6a–e).

Having found a strong causal relationship between acetate turnover and GSIS, we next examined the mechanism by which increased acetate turnover caused increased GSIS. We first investigated whether acetate could stimulate GSIS through a direct effect on  $\beta$ -cells, perhaps by increasing acetylcholine concentrations<sup>11,15</sup>. However, we found that neither acetate nor acetylcholine stimulated GSIS in isolated islet perfusions, ruling out a direct effect on  $\beta$ -cells (Fig. 3h and Extended Data Fig. 6f–h). In addition, concentrations of  $\beta$ -cell stimulatory amino acids and plasma glucagon were unchanged or reduced in the acetate-infused rats (Extended Data Fig. 6i–l). A small (about 2 pM) but significant ( $P < 0.05$ ) increase in plasma glucagon-like peptide-1 (GLP-1) concentration was measured in rats after 120 min of acetate infusion (Extended Data Fig. 6m). Because GLP-1 can stimulate GSIS<sup>12,13</sup>, we treated acetate-infused rats with a GLP-1 inhibitor; this treatment produced no change in GSIS (Extended Data Fig. 6n–s), demonstrating that these small changes in GLP-1 were not responsible for the increased GSIS in acetate-infused rats.

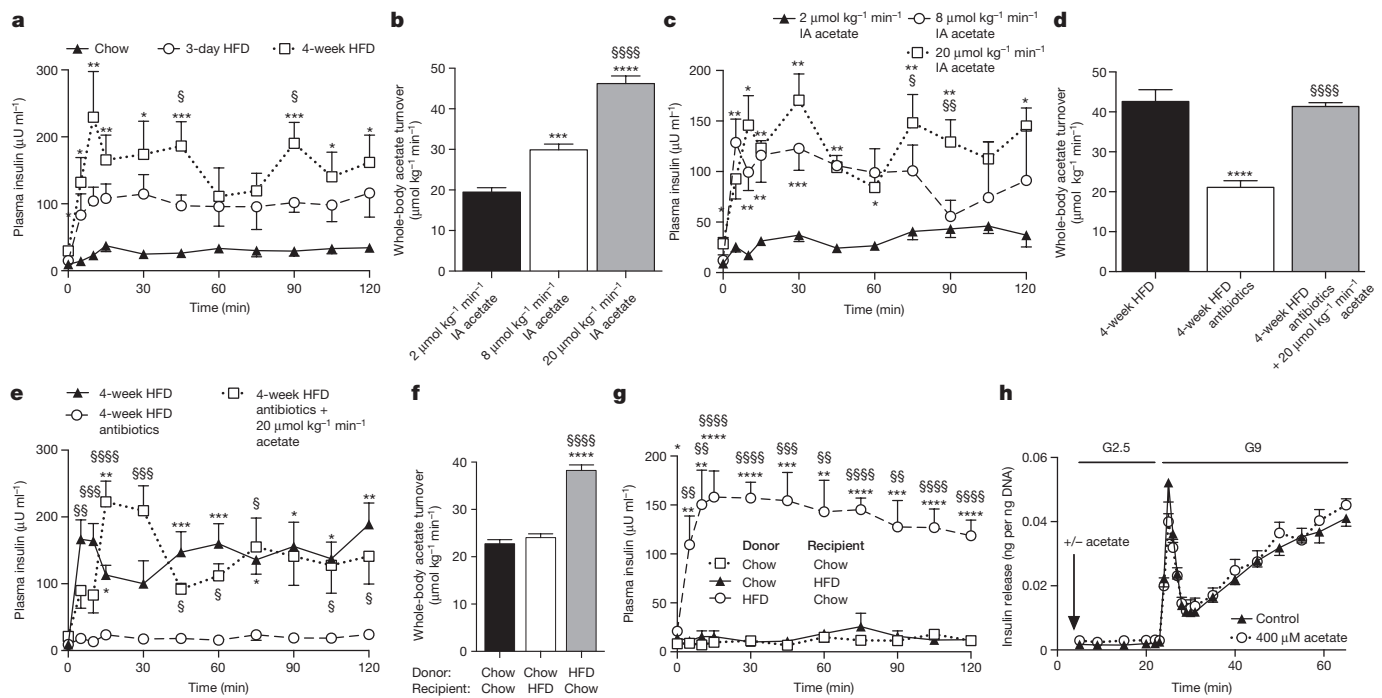
### Acetate drives GSIS via parasympathetic input

As parasympathetic input is a well-known stimulator of  $\beta$ -cell insulin secretion<sup>14</sup>, we next measured plasma gastrin concentrations as a marker of parasympathetic activity in rats acutely infused with acetate. Plasma gastrin increased threefold after 60 min of infusion with



**Figure 2 | The contents of the colonic lumen are the primary source of acetate in HFD-fed rats.** **a**, Tissue and plasma [<sup>13</sup>C]acetate enrichment in rats infused with [<sup>13</sup>C]acetate. **b**, Whole-body acetate turnover before and after washout of the gut.  $**P < 0.01$ ,  $***P < 0.001$  versus before washout. **c**, Whole-body acetate turnover in HFD-fed rats before and after portal vein ligation. **d, e**, Whole-body acetate turnover, and acetate in the entire caecum and colon lumen. **f**, Whole-body acetate turnover in HFD-fed rats before and after acute colectomy. In all panels,  $***P < 0.001$ ,  $****P < 0.0001$  by two-tailed unpaired Student's *t*-test. Data are the mean  $\pm$  s.e.m. of  $n = 6$  replicates per group.

20  $\mu\text{mol kg}^{-1} \text{ min}^{-1}$  acetate (Fig. 4a). Increases in brain acetate concentrations in the acetate-infused animals confirmed the ability of acetate infused systemically to enter the brain circulation (Fig. 4b). Because vagal stimulation has been shown to drive, and vagotomy has been shown to suppress, basal and glucose-stimulated insulin secretion<sup>16–20</sup>, we hypothesized that vagotomy would reduce GSIS in acetate-infused rats. Consistent with this hypothesis, vagotomized rats infused with acetate exhibited an approximately fourfold reduction in plasma insulin concentrations throughout a hyperglycaemic clamp without any change in plasma glucagon concentrations, when compared with intact rats infused with acetate (Fig. 4c and Extended Data Fig. 7a–h). In addition, treatment with the parasympathetic blocker atropine before the acetate infusion abolished the ability of acetate to stimulate GSIS, without any effect on plasma glucagon concentrations (Fig. 4d and Extended Data Fig. 7i–n), replicating prior studies demonstrating that atropine can suppress basal and glucose-stimulated insulin secretion indirectly *in vitro* and *in vivo*<sup>19,21–23</sup>. To test whether the effect of parasympathetic stimulation of GSIS is centrally mediated, we administered acetate by intracerebroventricular (ICV) injection at a dose chosen to increase cerebrospinal fluid acetate concentrations by 200  $\mu\text{M}$ , mimicking the increases in plasma acetate concentrations caused by intra-arterial infusion of 20  $\mu\text{mol kg}^{-1} \text{ min}^{-1}$  acetate. ICV acetate tripled GSIS during a hyperglycaemic clamp without inducing any difference in systemic acetate concentrations; however, this effect was blocked by treatment with atropine, and was independent of changes in plasma glucagon concentrations (Fig. 4e, f and Extended Data Fig. 8a–d), suggesting that acetate

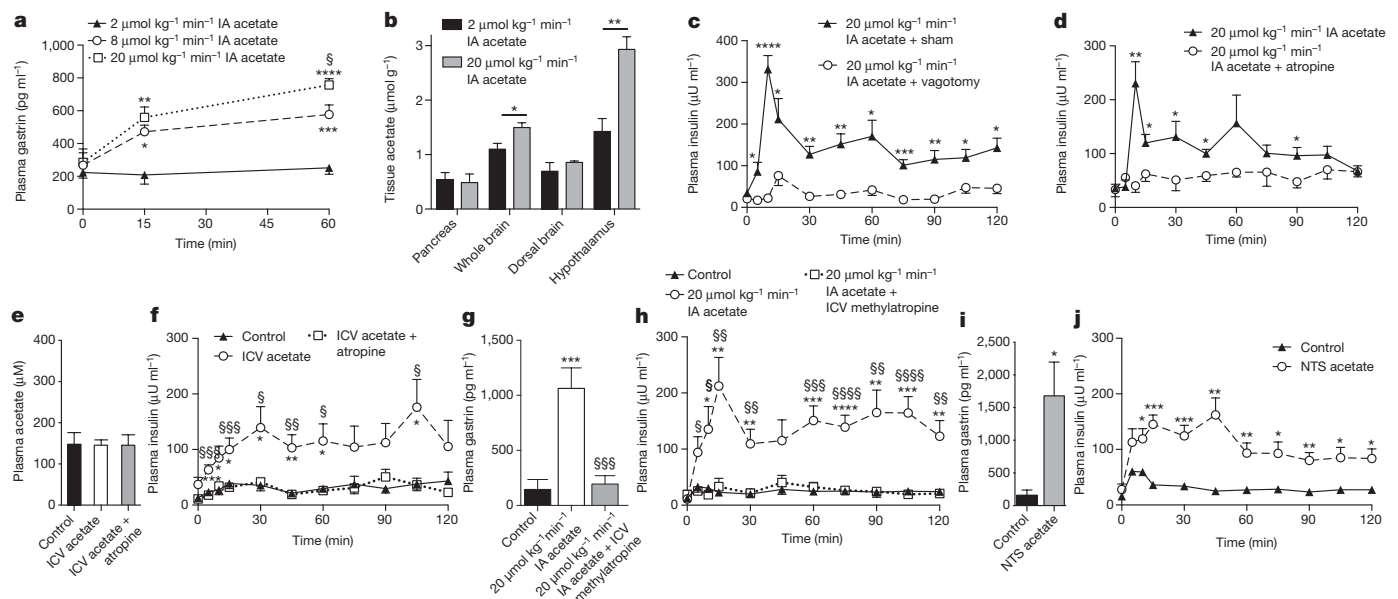


**Figure 3 | Acetate turnover drives GSIS.** **a**, Plasma insulin in a hyperglycaemic clamp.  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$  versus chow-fed rats;  $\$P < 0.05$  versus 3-day HFD-fed rats. **b, c**, Acetate turnover and GSIS in rats given acute acetate.  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.0001$  versus 2  $\mu\text{mol kg}^{-1} \text{min}^{-1}$ ;  $\$P < 0.05$ ,  $\$P < 0.01$ ,  $\$P < 0.0001$  versus 8  $\mu\text{mol kg}^{-1} \text{min}^{-1}$ . **d, e**, Acetate turnover and GSIS in rats treated with broad-spectrum, non-absorbable oral antibiotics.  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.0001$  versus controls;  $\$P < 0.05$ ,  $\$P < 0.01$ ,  $\$P < 0.001$ ,

$\$P < 0.0001$  versus antibiotic-treated rats. **f, g**, Whole-body acetate turnover and GSIS.  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ ,  $****P < 0.0001$  versus chow-fed donor and chow-fed recipient;  $\$P < 0.01$ ,  $\$P < 0.001$ ,  $\$P < 0.0001$  versus chow-fed donor and HFD-fed recipient. **h**, GSIS in isolated islets (KRB buffer;  $n = 4$  replicates per group). Data show mean  $\pm$  s.e.m. Groups were compared by one-way ANOVA with Bonferroni's multiple comparisons test (**a–g**) or by two-tailed unpaired Student's  $t$ -test (**h**). Unless otherwise specified,  $n = 6$  rats per group.

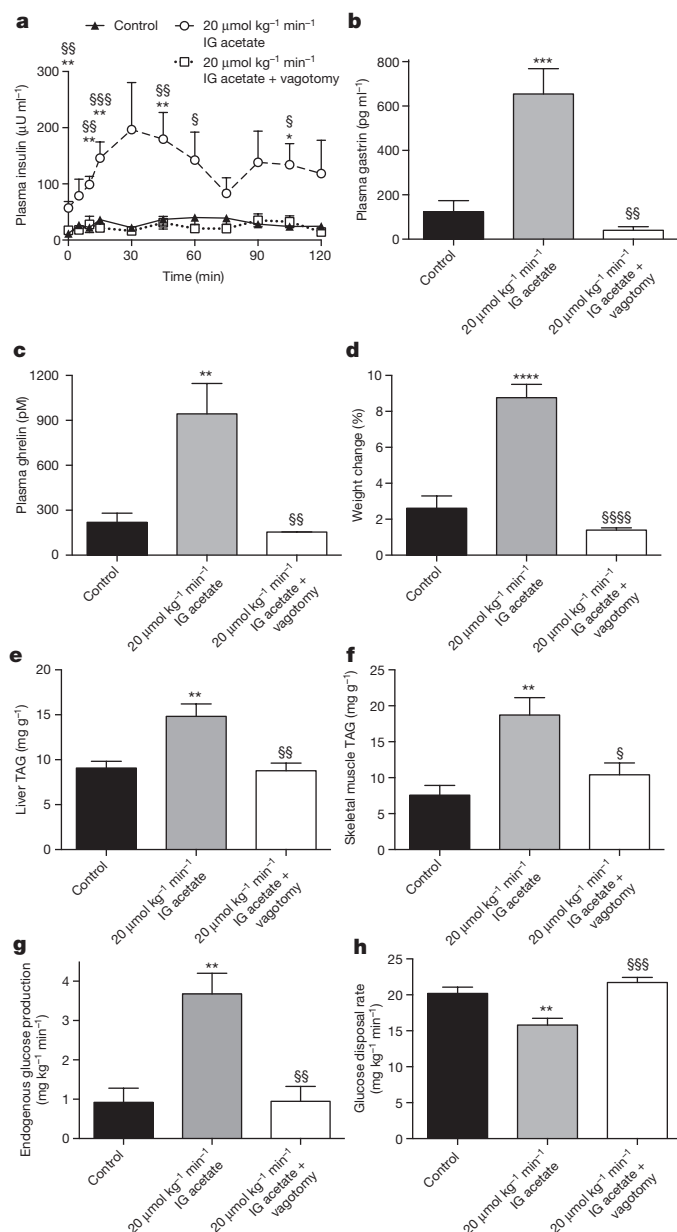
acts centrally to increase GSIS. Because atropine has also been shown to act directly on  $\beta$ -cells to suppress insulin secretion, we infused rats with intra-arterial acetate and examined the effect of ICV methylatropine,

an atropine analogue that does not cross the blood–brain barrier. Consistent with acetate driving GSIS via the parasympathetic nervous system, methylatropine fully abrogated the ability of acetate to drive



**Figure 4 | Acetate drives increased GSIS via parasympathetic activation.** **a**, Plasma gastrin.  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ ,  $****P < 0.0001$  versus 2  $\mu\text{mol kg}^{-1} \text{min}^{-1}$  acetate;  $\$P < 0.05$  versus 8  $\mu\text{mol kg}^{-1} \text{min}^{-1}$  acetate. **b**, Tissue acetate. **c, d**, GSIS. **e, f**, Plasma acetate and GSIS.  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$  versus controls;  $\$P < 0.05$ ,  $\$P < 0.01$ ,  $\$P < 0.001$  versus ICV acetate. **g, h**, Plasma gastrin (120 min) and GSIS.  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ ,  $****P < 0.0001$  versus controls;

$\$P < 0.05$ ,  $\$P < 0.01$ ,  $\$P < 0.001$ ,  $\$P < 0.0001$  versus acetate. **i, j**, Plasma gastrin (120 min) and GSIS following acetate injection into the nucleus tractus solitarius. Data are the mean  $\pm$  s.e.m. of  $n = 6$  animals per group, with groups compared by ANOVA with Bonferroni's multiple comparisons test (**a, e–h**) or by two-tailed unpaired Student's  $t$ -test (**b–d, i, j**). In **b–d, i**, and **j**,  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ ,  $****P < 0.0001$ .



**Figure 5 | Chronic increases in whole-body acetate turnover promote hyperphagia, obesity, and metabolic syndrome.** **a**, Plasma insulin during a hyperglycaemic clamp. **b**, **c**, Plasma gastrin and ghrelin at time 0 of the hyperglycaemic clamp. **d**, Weight change during the 10-day infusion ( $n = 16$  controls, 16 acetate, and 12 acetate plus vagotomy). **e**, **f**, Liver and skeletal muscle triglyceride content. **g**, Endogenous glucose production during a hyperinsulinaemic–euglycaemic clamp. **h**, Glucose disposal rate during the clamp. All data are mean  $\pm$  s.e.m. In all panels, \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$  versus controls; § $P < 0.05$ , §§ $P < 0.01$ , §§§ $P < 0.001$ , §§§§ $P < 0.0001$  versus acetate-treated rats by one-way ANOVA with Bonferroni's multiple comparisons test.  $n = 6$  replicates unless otherwise stated.

GSIS (Fig. 4g, h and Extended Data Fig. 8e–i). To confirm that activation of the parasympathetic nervous system drives this effect, we injected the same dose of acetate into the nucleus tractus solitarius, and found that this intervention replicated the effects of systemic and ICV acetate on GSIS by driving parasympathetic outflow as indicated by a tenfold increase in plasma gastrin concentrations without any change in systemic plasma acetate or glucagon (Fig. 4i, j and Extended Data Fig. 8j–o). Together, these data conclusively demonstrate that the acetate-induced increase in GSIS occurs through parasympathetic activation.

## Chronic increases in acetate drive obesity

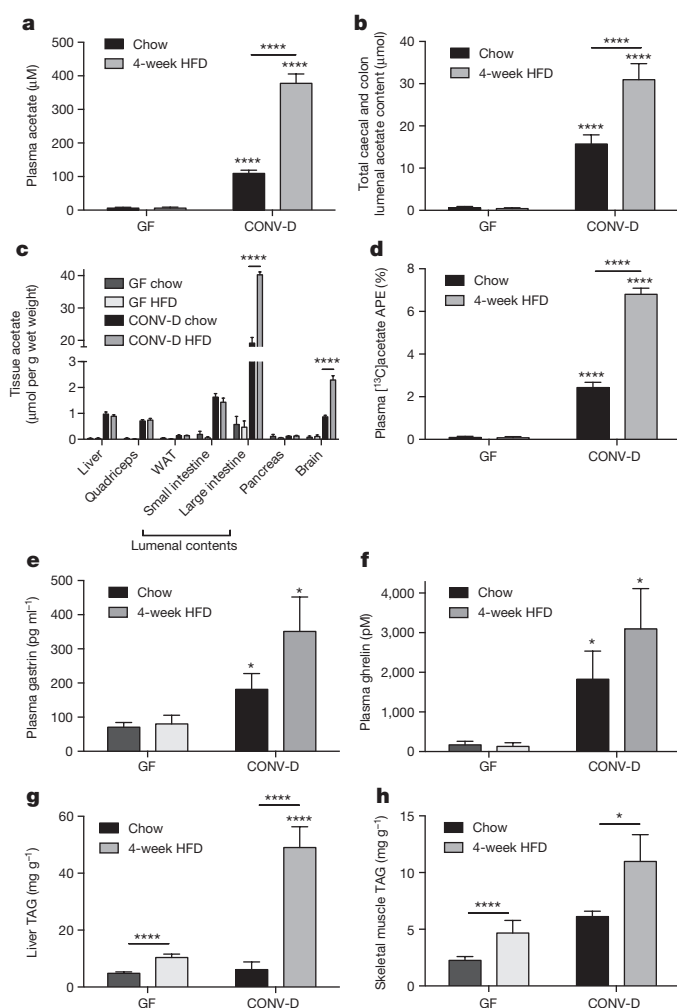
We next investigated whether a chronic increase in acetate turnover would promote chronic hyperinsulinaemia, hyperphagia, and weight gain and the associated sequelae of obesity. To answer this question, we performed continuous intragastric acetate infusions for 10 days, treating chow-fed rats with  $20 \mu\text{mol kg}^{-1} \text{min}^{-1}$  acetate to mimic the increase in gut microbial acetate production measured in HFD-fed rats (Extended Data Fig. 9a, b). Rats that received chronic intragastric acetate infusions exhibited increased insulin secretion during both a hyperglycaemic clamp and an intraperitoneal glucose tolerance test; this increase in insulin secretion was associated with a fivefold increase in plasma gastrin concentration (Fig. 5a, b and Extended Data Fig. 9c–h). All of these effects were prevented by vagotomy. Consistent with the hypothesis that chronic postprandial hyperinsulinaemia leads to increased weight gain, rats that received chronic intragastric acetate infusions exhibited more than a doubling in daily caloric intake and in weight gain over the ten-day infusion, which may be attributable, at least in part, to a threefold increase in plasma ghrelin concentrations (Fig. 5c, d and Extended Data Fig. 9i, j). These effects were also prevented by vagotomy, demonstrating that parasympathetic activation is necessary to mediate the effects of chronic acetate on GSIS in awake, unrestrained rats. Finally, acetate-infused rats exhibited insulin resistance, as indicated by impaired glucose disposal and impaired insulin suppression of hepatic glucose production during a hyperinsulinaemic–euglycaemic clamp, and increases in plasma, liver, and skeletal muscle triglyceride content without any changes in plasma glucagon concentrations (Fig. 5e–h and Extended Data Fig. 9k–p). Vagotomized rats exhibited none of these consequences of acetate infusion.

Together, these findings strongly suggest that the gut microbiota are responsible for generating increased acetate turnover and driving obesity in HFD-fed rats, although we cannot rule out the possibility that the microbiota also modulate acetate absorption<sup>24</sup>. To conclusively test the hypothesis that the gut microbiota are primarily responsible for increasing acetate turnover in HFD-fed rodents, we measured plasma and colonic acetate content in germ-free mice lacking gut microbes and ex-germ-free mice 4 weeks after colonization with normal mouse faeces (conventionalized; CONV-D) fed either a regular chow diet or HFD. Demonstrating the role of the gut microbiota as the main producer of acetate *in vivo*, germ-free mice had negligible plasma, colonic lumen, and tissue acetate concentrations as compared to CONV-D mice; only conventionalized mice exhibited an increase in acetate concentrations on HFD (Fig. 6a–c). Germ-free mice fed [<sup>13</sup>C]bicarbonate also exhibited strikingly lower plasma and tissue <sup>13</sup>C enrichment than CONV-D mice. Furthermore, in CONV-D mice but not germ-free mice, [<sup>13</sup>C] acetate was doubled in HFD-fed relative to chow-fed animals (Fig. 6d). As rodents do not possess the enzymes necessary to convert bicarbonate to acetate, these results demonstrate that the gut microbiota are responsible for the increased acetate turnover in HFD-fed animals, a conclusion corroborated by the fact that colonic [<sup>13</sup>C]acetate enrichment in CONV-D mice was more than double the enrichment in plasma or in any other tissue (Extended Data Fig. 10a). In contrast, propionate and butyrate concentration and enrichment were minimal in plasma and tissues in all mice (Extended Data Fig. 10b–g). Finally, because of the role of increased acetate in promoting parasympathetic activation, we measured plasma gastrin and ghrelin concentrations in the germ-free and CONV-D mice, and found that CONV-D mice exhibited two- and tenfold increases in gastrin and ghrelin, respectively, which were associated with two- and fivefold increases in liver and skeletal muscle triglyceride content (Fig. 6e–h), compared with the germ-free mice. Together, these data clearly implicate the gut microbiota as being responsible for the majority of the whole-body plasma acetate turnover *in vivo* and for the increase in acetate turnover observed in HFD-fed rats.

## Conclusions

In summary, we show here that increased acetate production due to a gut microbiota–nutrient interaction in HFD-fed rodents leads





**Figure 6 | Gut bacteria are responsible for the majority of acetate production *in vivo*, and for the increase in HFD-fed rodents.**

**a, b**, Plasma and caecal/colon lumen acetate concentrations in germ-free (GF) and CONV-D mice. **c**, Tissue acetate concentrations. **d**, Plasma [ $^{13}\text{C}$ ]acetate enrichment in mice fed water containing [ $^{13}\text{C}$ ]bicarbonate for 3 days. **e, f**, Plasma gastrin and ghrelin concentrations. **g, h**, Liver and skeletal muscle triglyceride content. In all panels, \* $P < 0.05$  and \*\*\*\* $P < 0.0001$ . Data are the mean  $\pm$  s.e.m. of  $n = 9$  GF mice and 10 CONV-D mice per diet.

to activation of the parasympathetic nervous system and results in increased ghrelin secretion and GSIS. This generates a positive feedback loop, resulting in hyperphagia, hypertriglyceridaemia, ectopic lipid deposition in liver and skeletal muscle, and liver and muscle insulin resistance (Extended Data Fig. 1). The increased acetate production that occurs when the gut microbiota are exposed to calorically dense nutrients may mediate an important positive feedback loop between the gut microbiota and the CNS that promotes hyperphagia (due to increased ghrelin secretion) and increased energy storage as fat (due to increased GSIS) in foraging animals when they stumble across calorically dense foodstuffs in the wilderness. However, in the setting of chronic exposure to calorically dense, abundant food, this gut microbiota–brain– $\beta$ -cell axis promotes obesity and its related sequelae of hyperlipidaemia, fatty liver disease and insulin resistance.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 12 May 2015; accepted 11 May 2016.

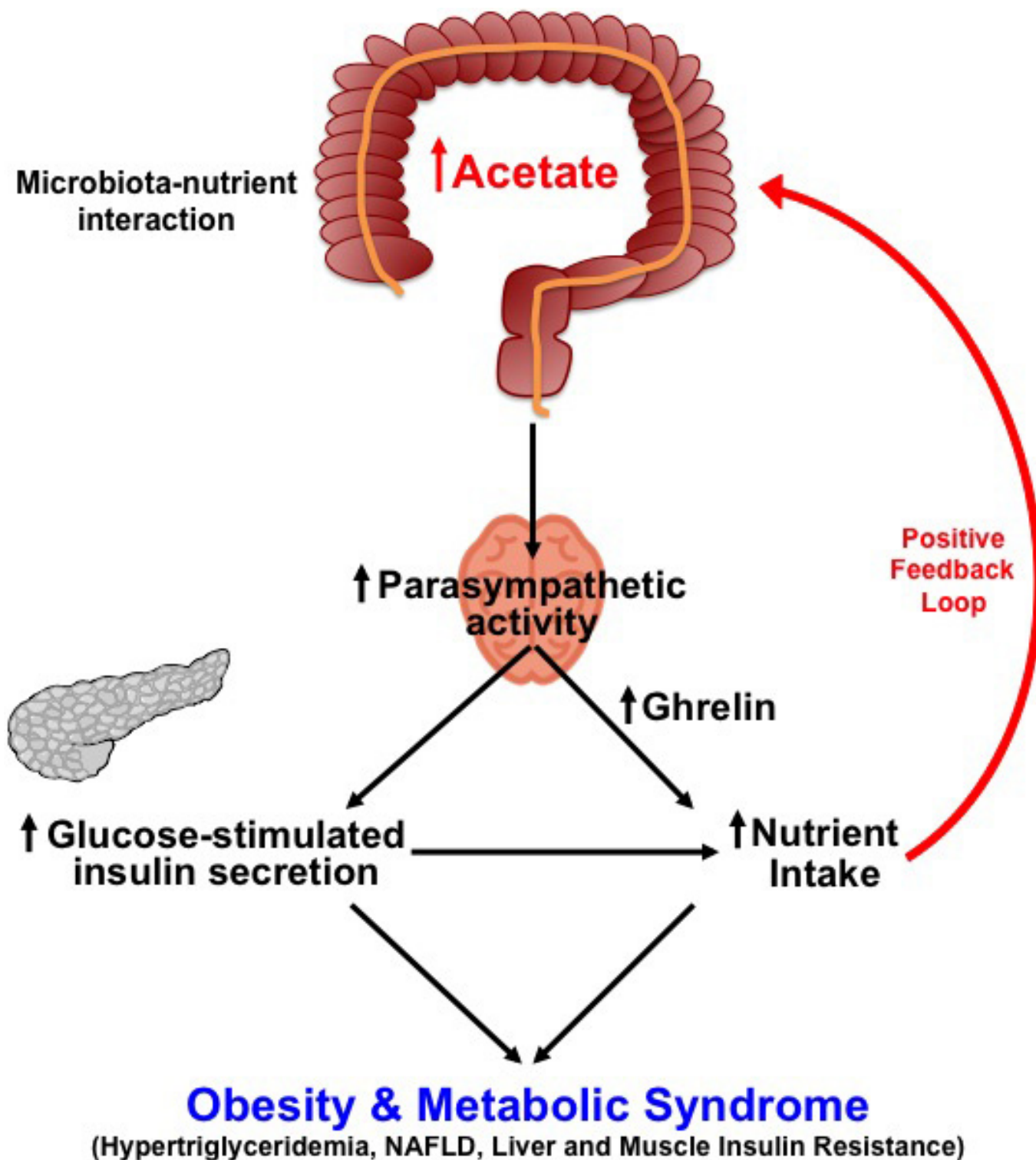
- Rahat-Rozenbloom, S., Fernandes, J., Gloor, G. B. & Wolever, T. M. Evidence for greater production of colonic short-chain fatty acids in overweight than lean humans. *Int. J. Obes.* **38**, 1525–1531 (2014).
- Shepherd, M. L., Ponder, M. A., Burk, A. O., Milton, S. C. & Swecker, W. S. Jr. Fibre digestibility, abundance of faecal bacteria and plasma acetate concentrations in overweight adult mares. *J. Nutr. Sci.* **3**, e10 (2014).
- Turnbaugh, P. J. *et al.* An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**, 1027–1031 (2006).
- Fernandes, J., Su, W., Rahat-Rozenbloom, S., Wolever, T. M. & Comelli, E. M. Adiposity, gut microbiota and faecal short chain fatty acids are linked in adult humans. *Nutr. Diabetes* **4**, e121 (2014).
- Li, M. *et al.* Gut carbohydrate metabolism instead of fat metabolism regulated by gut microbes mediates high-fat diet-induced obesity. *Benef. Microbes* **5**, 335–344 (2014).
- Murugesan, S. *et al.* Study of the diversity and short-chain fatty acids production by the bacterial community in overweight and obese Mexican children. *Eur. J. Clin. Microbiol. Inf. Dis.* **34**, 1337–1346 (2015).
- Murphy, E. F. *et al.* Composition and energy harvesting capacity of the gut microbiota: relationship to diet, obesity and time in mouse models. *Gut* **59**, 1635–1642 (2010).
- David, L. A. *et al.* Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**, 559–563 (2014).
- La-ongkham, O., Nakphachit, M., Leelavatcharamas, V., Keawsonpong, S. & Nitisinprasert, S. Distinct gut microbiota of healthy children from two different geographic regions of Thailand. *Arch. Microbiol.* **197**, 561–573 (2015).
- Aguirre, M., Jonkers, D. M., Troost, F. J., Roeselers, G. & Venema, K. *In vitro* characterization of the impact of different substrates on metabolite production, energy extraction and composition of gut microbiota from lean and obese subjects. *PLoS One* **9**, e113864 (2014).
- Shen, W. *et al.* Protective effects of *R*-alpha-lipoic acid and acetyl-L-carnitine in MIN6 and isolated rat islet cells chronically exposed to oleic acid. *J. Cell. Biochem.* **104**, 1232–1243 (2008).
- Drucker, D. J. Minireview: the glucagon-like peptides. *Endocrinology* **142**, 521–527 (2001).
- MacDonald, P. E. *et al.* The multiple actions of GLP-1 on the process of glucose-stimulated insulin secretion. *Diabetes* **51** (suppl. 3), S434–S442 (2002).
- Ahrén, B. Autonomic regulation of islet hormone secretion—implications for health and disease. *Diabetologia* **43**, 393–410 (2000).
- Ronnebaum, S. M. *et al.* Chronic suppression of acetyl-CoA carboxylase 1 in  $\beta$ -cells impairs insulin secretion via inhibition of glucose rather than lipid metabolism. *J. Biol. Chem.* **283**, 14248–14256 (2008).
- Ionescu, E., Rohner-Jeanrenaud, F., Berthoud, H. R. & Jeanrenaud, B. Increases in plasma insulin levels in response to electrical stimulation of the dorsal motor nucleus of the vagus nerve. *Endocrinology* **112**, 904–910 (1983).
- Sakaguchi, T. & Yamaguchi, K. Effects of electrical stimulation of the hepatic vagus nerve on the plasma insulin concentration in the rat. *Brain Res.* **164**, 314–316 (1979).
- Lee, K. C. & Miller, R. E. The hepatic vagus nerve and the neural regulation of insulin secretion. *Endocrinology* **117**, 307–314 (1985).
- Frohman, L. A., Ezdinli, E. Z. & Javid, R. Effect of vagotomy and vagal stimulation on insulin secretion. *Diabetes* **16**, 443–448 (1967).
- Bergman, R. N. & Miller, R. E. Direct enhancement of insulin secretion by vagal stimulation of the isolated pancreas. *Am. J. Physiol.* **225**, 481–486 (1973).
- Ahrén, B. & Holst, J. J. The cephalic insulin response to meal ingestion in humans is dependent on both cholinergic and noncholinergic mechanisms and is important for postprandial glycemia. *Diabetes* **50**, 1030–1038 (2001).
- Yamazaki, H., Philbrick, W., Zawulich, K. C. & Zawulich, W. S. Acute and chronic effects of glucose and carbachol on insulin secretion and phospholipase C activation: studies with diazoxide and atropine. *Am. J. Physiol. Endocrinol. Metab.* **290**, E26–E33 (2006).
- D'Alessio, D. A., Kieffer, T. J., Taborsky, G. J., Jr & Havel, P. J. Activation of the parasympathetic nervous system is necessary for normal meal-induced insulin secretion in rhesus macaques. *J. Clin. Endocrinol. Metab.* **86**, 1253–1259 (2001).
- Wichmann, A. *et al.* Microbial modulation of energy availability in the colon regulates intestinal transit. *Cell Host Microbe* **14**, 582–590 (2013).

**Supplementary Information** is available in the online version of the paper.

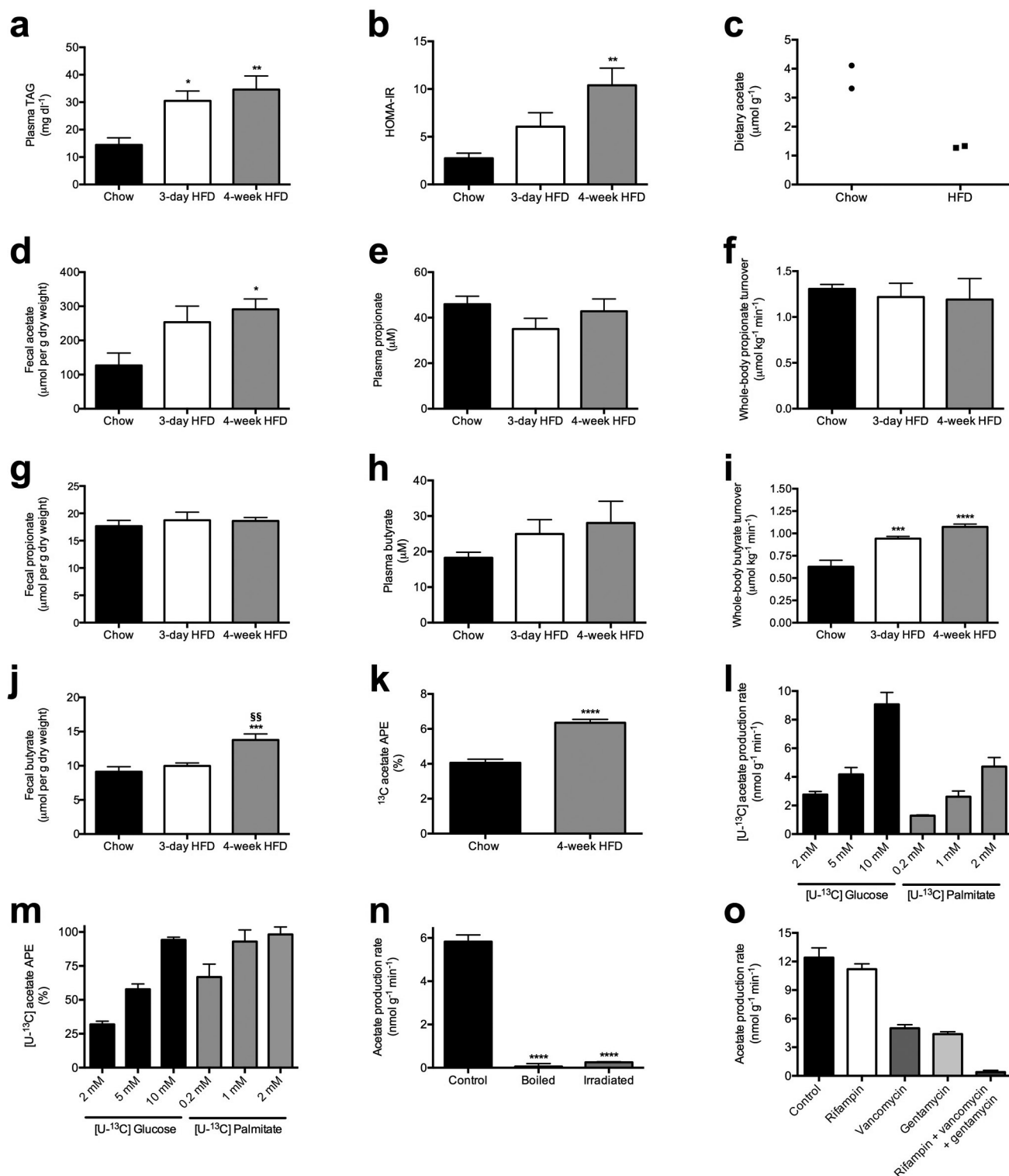
**Acknowledgements** We thank J. Dong, Y. Nozaki, W. Zhu, X. Zhao, J. Stack, M. Kahn, B. Lim, and Y. Kosover for technical assistance. This study was funded by grants from the National Institutes of Health (R01 DK-40936, R01 AG-23686, P30 DK-45735, U24 DK-59635, T32 DK-101019, R01 DK-92606, R01 GM-103574 and DP2 GM-105456) and the Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen.

**Author Contributions** R.J.P., L.P., N.A.B., G.W.C., D.Z., and R.L.C. performed the *in vivo* and *in vitro* studies, and all authors analysed data. G.W.C., K.F.P., R.G.K., A.L.G., and G.I.S. provided critical advice for the experiments. Studies were designed and the manuscript was written by R.J.P. and G.I.S. with input from all authors.

**Author Information** Sequence data are deposited to the European Nucleotide Archive with accession code PRJEB13505. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.I.S. ([gerald.shulman@yale.edu](mailto:gerald.shulman@yale.edu)).



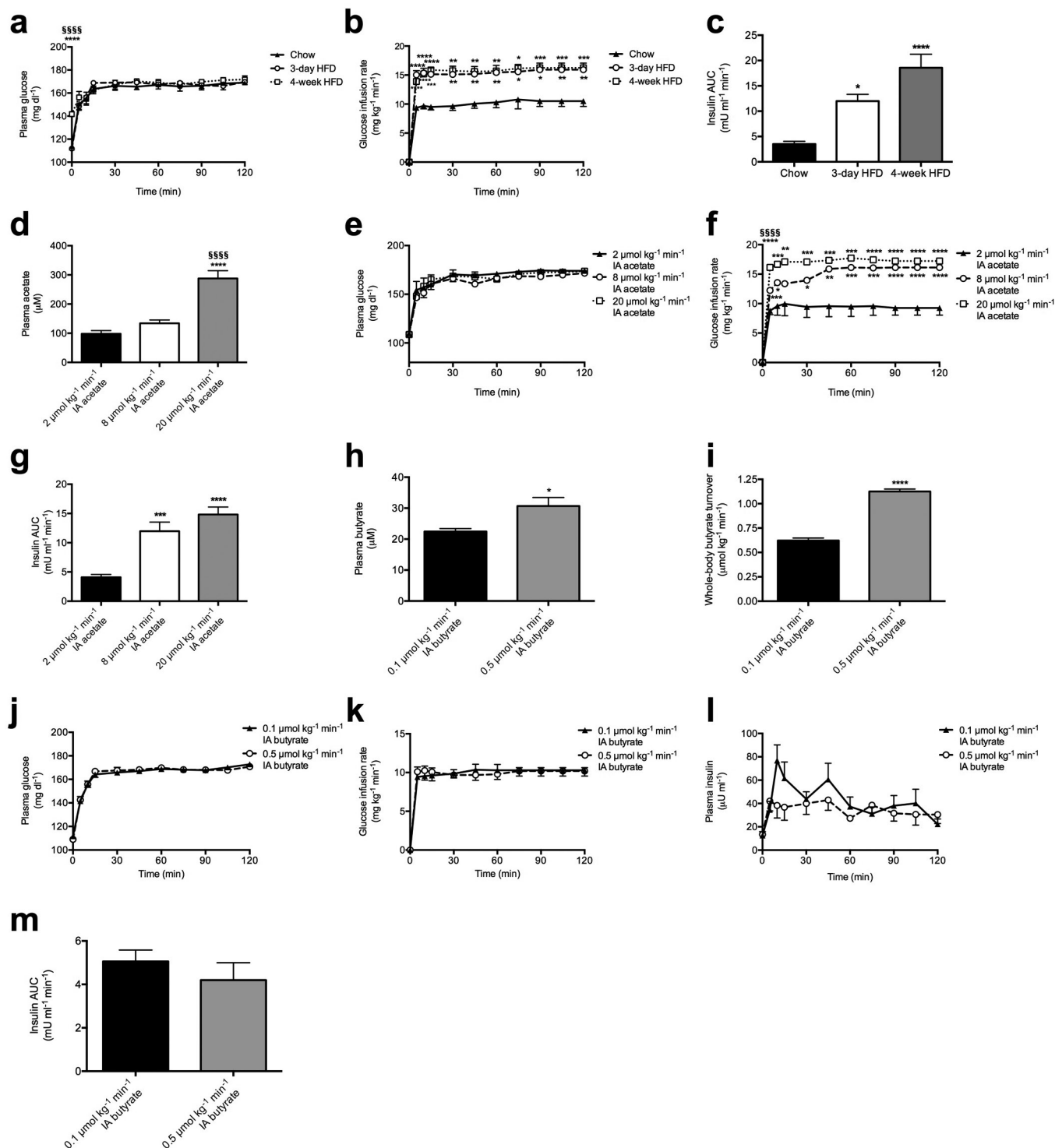
Extended Data Figure 1 | Mechanism by which a diet-microbiota interaction drives obesity and the metabolic syndrome.



**Extended Data Figure 2 | HFD-fed rats exhibit increased gut acetate production.** **a**, Plasma triglycerides. **b**, HOMA-IR. **c**, Dietary acetate concentrations.  $n = 2$  replicates per diet. **d**, Faecal acetate normalized to dry weight. **e–g**, Plasma propionate, whole-body propionate turnover, and faecal propionate concentrations. **h–j**, Plasma butyrate, whole-body butyrate turnover, and faecal butyrate concentrations. **k**, [<sup>13</sup>C]acetate enrichment in plasma of rats fed [<sup>13</sup>C]bicarbonate-labelled food and water. **l**, [U-<sup>13</sup>C]acetate from faeces incubated in [U-<sup>13</sup>C]glucose or fatty acids. **m**, *In vitro* acetate production rate from faeces incubated in [U-<sup>13</sup>C] glucose or fatty acids. **n**, *In vitro* acetate production rate in control, boiled,

and UV-irradiated faecal samples. \*\*\*\* $P < 0.0001$  versus control. **o**, *In vitro* faecal acetate production following treatment with antibiotics. Unless otherwise specified, \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$  versus chow-fed rats; §§ $P < 0.01$  versus 3-day HFD-fed rats by one-way ANOVA with Bonferroni's multiple comparisons test. In **k**, data were compared by two-tailed unpaired Student's *t*-test. In **l–o**, data are the mean  $\pm$  s.e.m. of  $n = 4$  per group, with comparisons to controls via two-tailed unpaired Student's *t*-test (**n**). Unless otherwise specified,  $n = 6$  replicates per group.

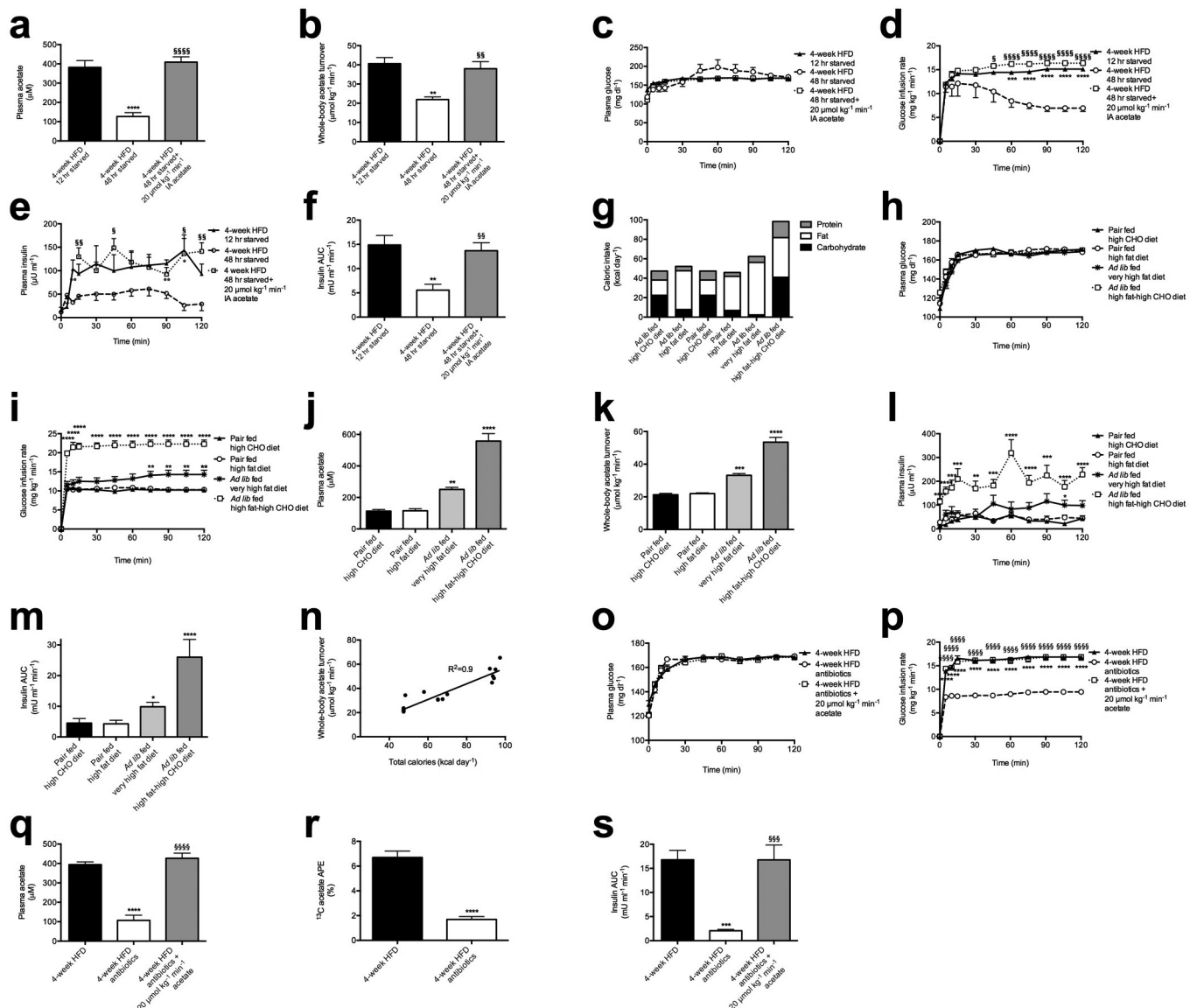




### Extended Data Figure 3 | HFD-fed rats exhibit increased GSIS

driven by increased acetate turnover. **a**, **b**, Plasma glucose and glucose infusion rate during a hyperglycaemic clamp. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$  versus chow-fed rats. **c**, Plasma insulin area under the curve (AUC) during the hyperglycaemic clamp. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$  versus 2 μmol kg<sup>-1</sup> min<sup>-1</sup> acetate; \$\$\$\$ $P < 0.0001$  versus 8 μmol kg<sup>-1</sup> min<sup>-1</sup> acetate. **e**, **f**, Plasma glucose and glucose infusion rate during a

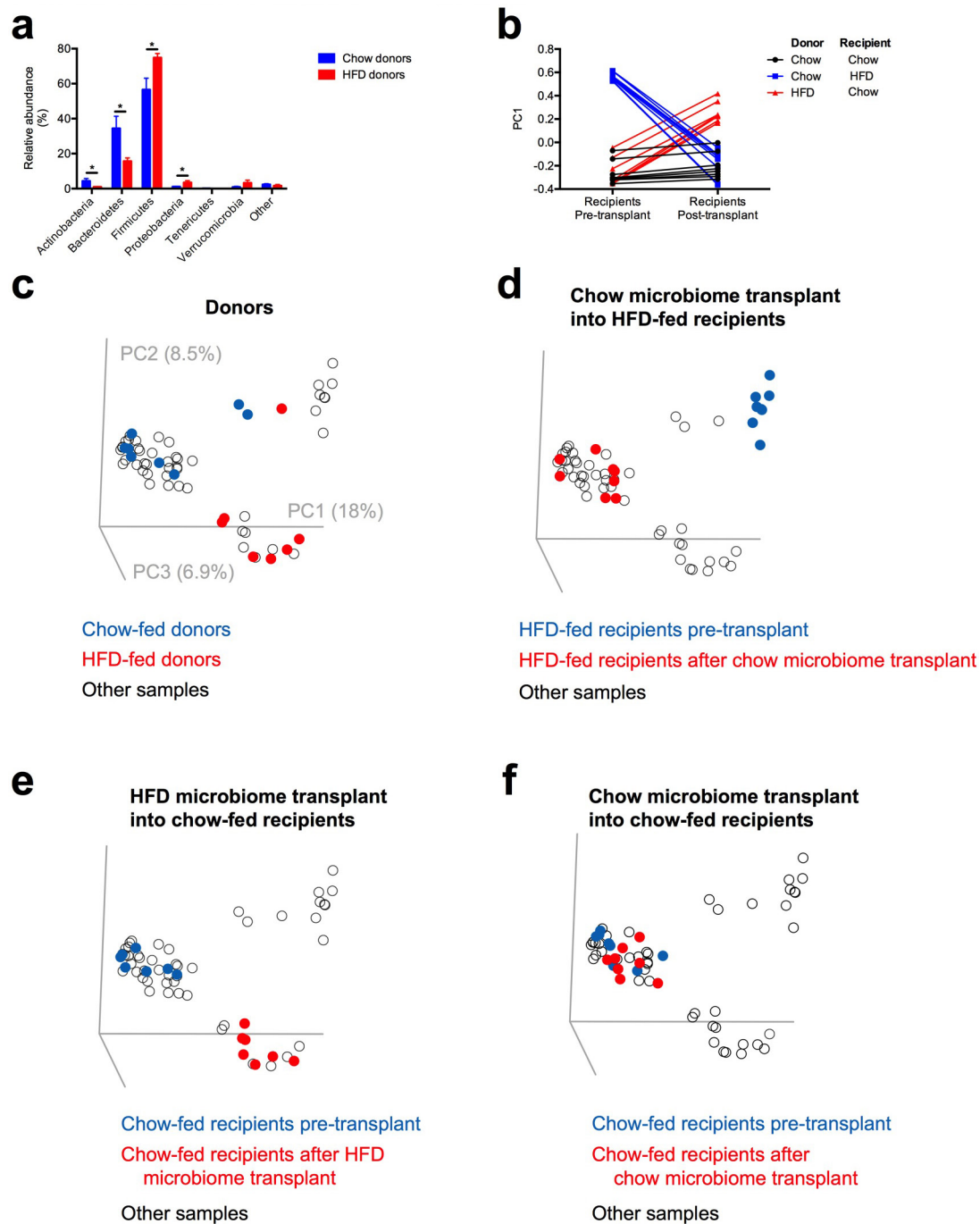
hyperglycaemic clamp. **g**, Plasma insulin AUC during the hyperglycaemic clamp. **h**, **i**, Plasma butyrate and whole-body butyrate turnover. \* $P < 0.05$ , \*\*\*\* $P < 0.0001$ . **j**, **k**, Plasma glucose and glucose infusion rate during a hyperglycaemic clamp. **l**, **m**, Plasma insulin concentrations during the hyperglycaemic clamp, and plasma insulin AUC. In all panels, data are the mean ± s.e.m. of  $n = 6$  animals per group, with comparisons by one-way ANOVA with Bonferroni's multiple comparisons test (**a–g**) or two-tailed unpaired Student's  $t$ -test (**h–m**).



**Extended Data Figure 4 | Increasing total caloric intake leads to increased acetate turnover and GSIS via the microbiota in rats.**

**a, b**, Plasma acetate and whole-body acetate turnover. **c, d**, Plasma glucose and glucose infusion rate during a hyperglycaemic clamp. **e, f**, Plasma insulin and insulin AUC during the clamp. **g**, Caloric intake from protein, fat, and carbohydrate. In **g–m**, each group was compared to pair-fed, high-carbohydrate-fed rats. **h, i**, Plasma glucose and glucose infusion rate in the hyperglycaemic clamp. **j, k**, Plasma acetate and whole-body acetate turnover. **l, m**, Plasma insulin and insulin AUC during the hyperglycaemic clamp. **n**, Linear regression: whole-body acetate turnover versus total caloric intake in each diet group. **o, p**, Plasma glucose and glucose infusion rate during a hyperglycaemic clamp in 4-week HFD-fed rats treated with

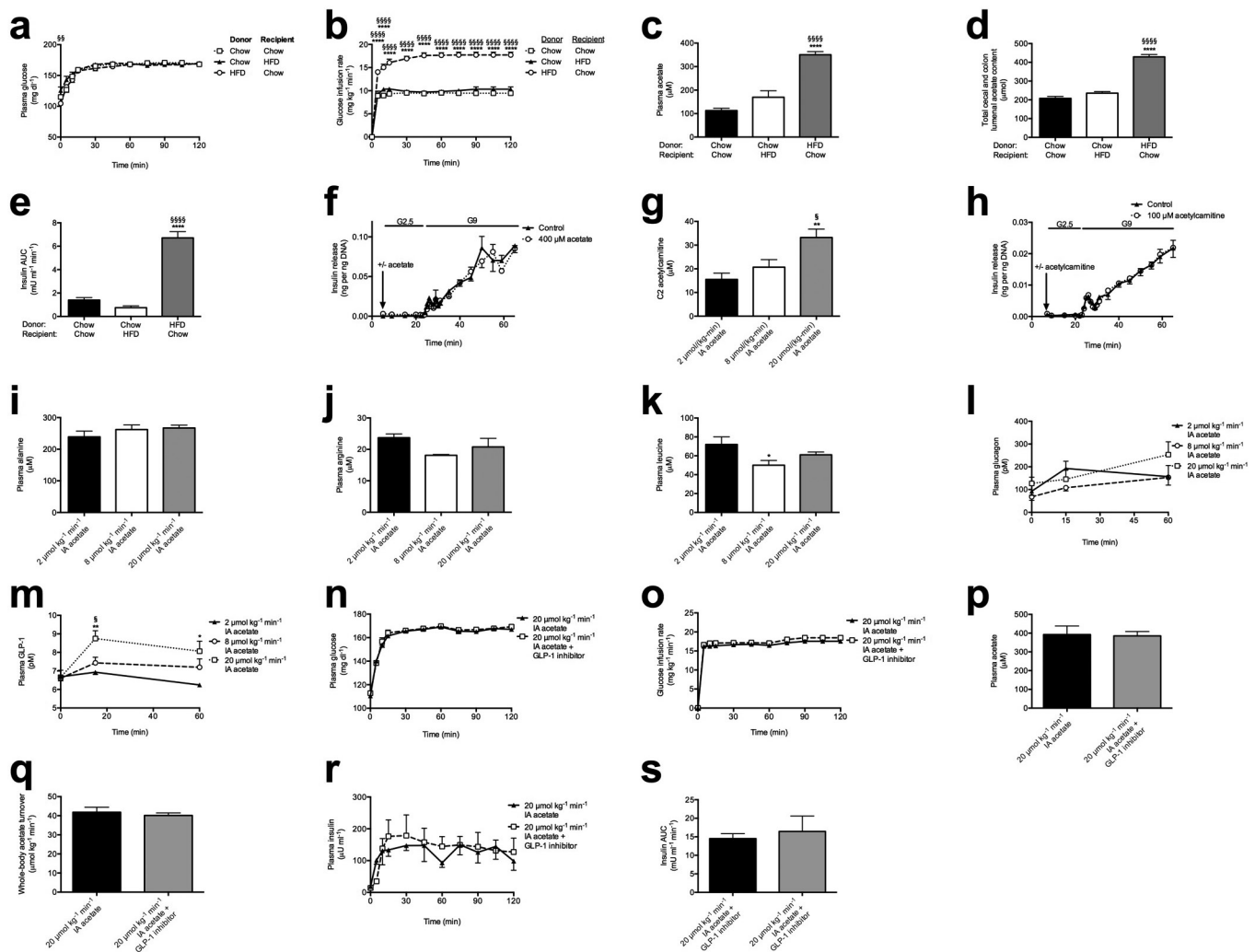
broad-spectrum non-absorbable antibiotics. **q**, Plasma acetate. **r**, Plasma [ $^{13}\text{C}$ ]acetate enrichment following three days of feeding [ $^{13}\text{C}$ ]bicarbonate food and water. Data were compared using the two-tailed unpaired Student's *t*-test. **s**, Insulin AUC during a hyperglycaemic clamp. In all panels, data are the mean  $\pm$  s.e.m. of  $n = 6$  rats per group, with groups compared by one-way ANOVA with Bonferroni's multiple comparisons test, unless otherwise stated. In **a–f**,  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ ,  $****P < 0.0001$  versus 12-h starved rats;  $\$P < 0.05$ ,  $$$$P < 0.01$ ,  $$$$$P < 0.001$  versus 48-h starved rats. In **h–m**,  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ ,  $****P < 0.0001$  versus pair-fed rats given the high-carbohydrate diet. In **o–s**,  $***P < 0.001$ ,  $****P < 0.0001$  versus HFD-fed rats;  $$$$P < 0.001$ ,  $$$$$P < 0.0001$  versus antibiotics-treated rats.



**Extended Data Figure 5 | Faecal transplantation alters recipient microbiomes to resemble their donors as revealed by culture-independent 16S rRNA sequencing of faecal microbiomes from donors and recipients.** **a**, Relative abundance at the phylum level. Only phyla with relative abundance  $\geq 0.1\%$  in at least one group are shown. Data are the mean  $\pm$  s.e.m. of  $n = 7-8$  replicates per group;  $*P < 0.05$  by 2-tailed unpaired Student's  $t$ -test. **b-f**, Beta diversity analysis of faecal microbiomes

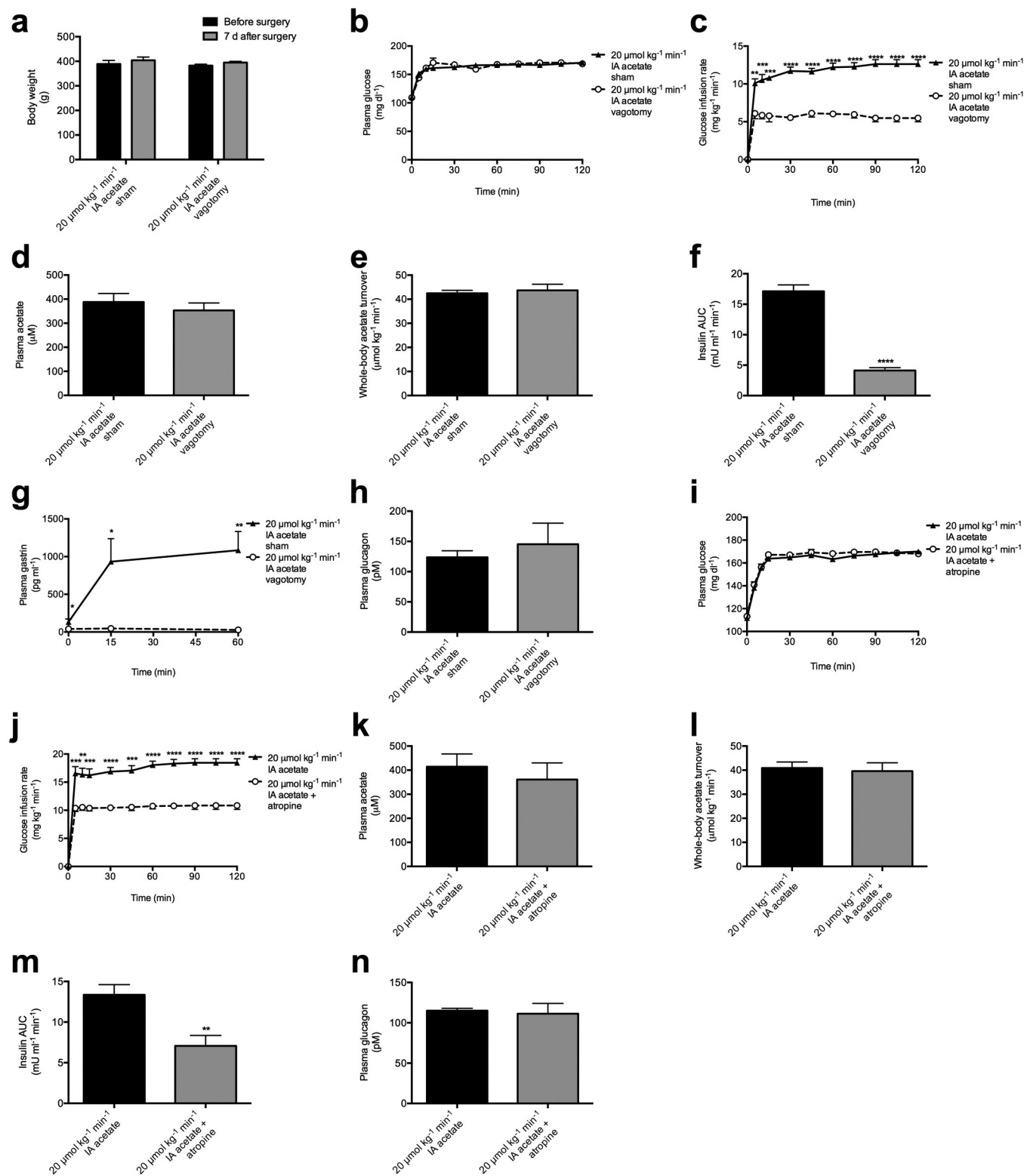
before and after transplantation. The largest component of variation (captured by principal coordinate (PC)1) is shown in **b** and PC1–PC3 are shown in **c–f**. Rats from independent litters were randomized before diet administration or faecal transplantation. Beta diversity reflects principal coordinates analysis based on Hellinger distances; the results from unweighted, non-phylogenetic distance metrics and from phylogenetic metrics (weighted and unweighted UniFrac) are similar.





**Extended Data Figure 6 | The gut microbiota drive increased acetate turnover and GSIS.** **a, b**, Plasma glucose and glucose infusion rate during a hyperglycaemic clamp in rats following faecal transplant replicates acetate turnover and GSIS in the donor group. **c**, Plasma acetate. **d**, Faecal acetate concentration.  $n = 7$  (HFD to chow) or 8 (chow to chow, chow to HFD) per group. **e**, Plasma insulin AUC. **f**, Glucose-stimulated insulin release in isolated islets incubated with 400 μM acetate in a physiological buffer.  $n = 4$  per group. **g**, Plasma C2 acetylcarnitine content. **h**, Glucose-stimulated insulin release in isolated islets incubated with 100 μM acetylcarnitine. **i–m**, Plasma alanine, leucine, arginine, glucagon, and GLP-1 concentrations. **n, o**, Plasma glucose and glucose infusion rate during a hyperglycaemic clamp in acetate-infused rats treated with a

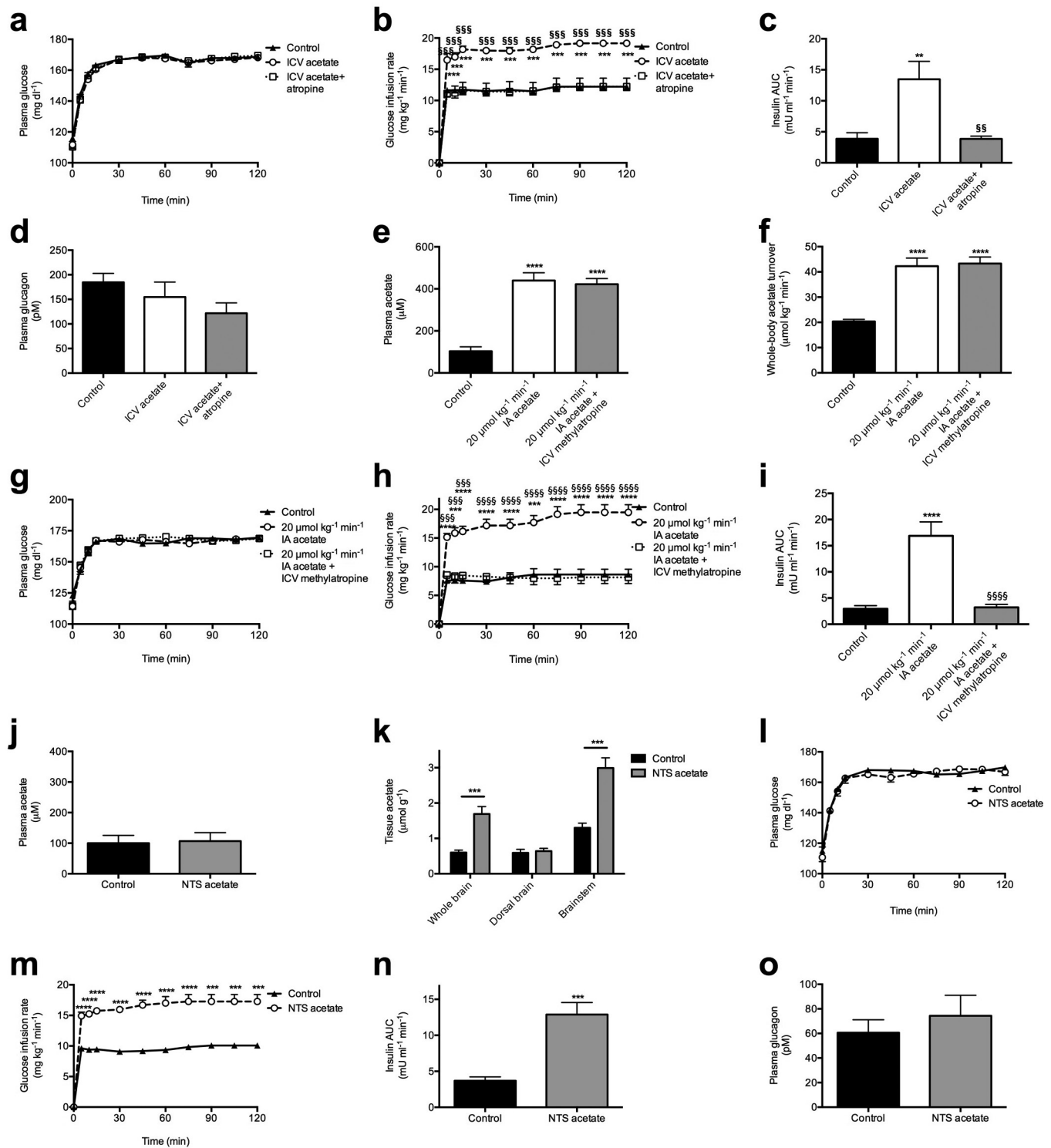
GLP-1 inhibitor. **p, q**, Plasma acetate and whole-body acetate turnover. **r, s**, Plasma insulin and insulin AUC during the clamp. In all panels, data are mean  $\pm$  s.e.m. of  $n = 6$  per group. In **a–e**, \*\*\*\* $P < 0.0001$  versus chow-fed donor to chow-fed recipient transplants; \*\*\*\* $P < 0.0001$  versus chow-fed donor to HFD-fed recipient transplants by one-way ANOVA with Bonferroni's multiple comparisons test. Data are the mean  $\pm$  s.e.m. of  $n = 6$  (HFD to chow) or 7 (chow to chow, chow to HFD) per group. In **g–m**, \* $P < 0.05$ , \*\* $P < 0.01$  versus 2 μmol kg<sup>-1</sup> min<sup>-1</sup> acetate; \$ $P < 0.05$  versus 8 μmol kg<sup>-1</sup> min<sup>-1</sup> acetate by one-way ANOVA with Bonferroni's multiple comparisons test. In **g–s**, data are the mean  $\pm$  s.e.m. of  $n = 6$  (unless otherwise specified) per group. In **n–s**, no significant differences were measured by the two-tailed unpaired Student's  $t$ -test.



### Extended Data Figure 7 | Acetate drives GSIS via a CNS mechanism.

**a**, Body weight before and after vagotomy. **b**, **c**, Plasma glucose and glucose infusion rate during a hyperglycaemic clamp. **d**, **e**, Plasma acetate and whole-body acetate turnover. **f**, Insulin AUC during the clamp. **g**, Plasma gastrin during the clamp. **h**, Plasma glucagon after 120 min of the clamp. **i**, **j**, Plasma glucose and glucose infusion rate during a hyperglycaemic

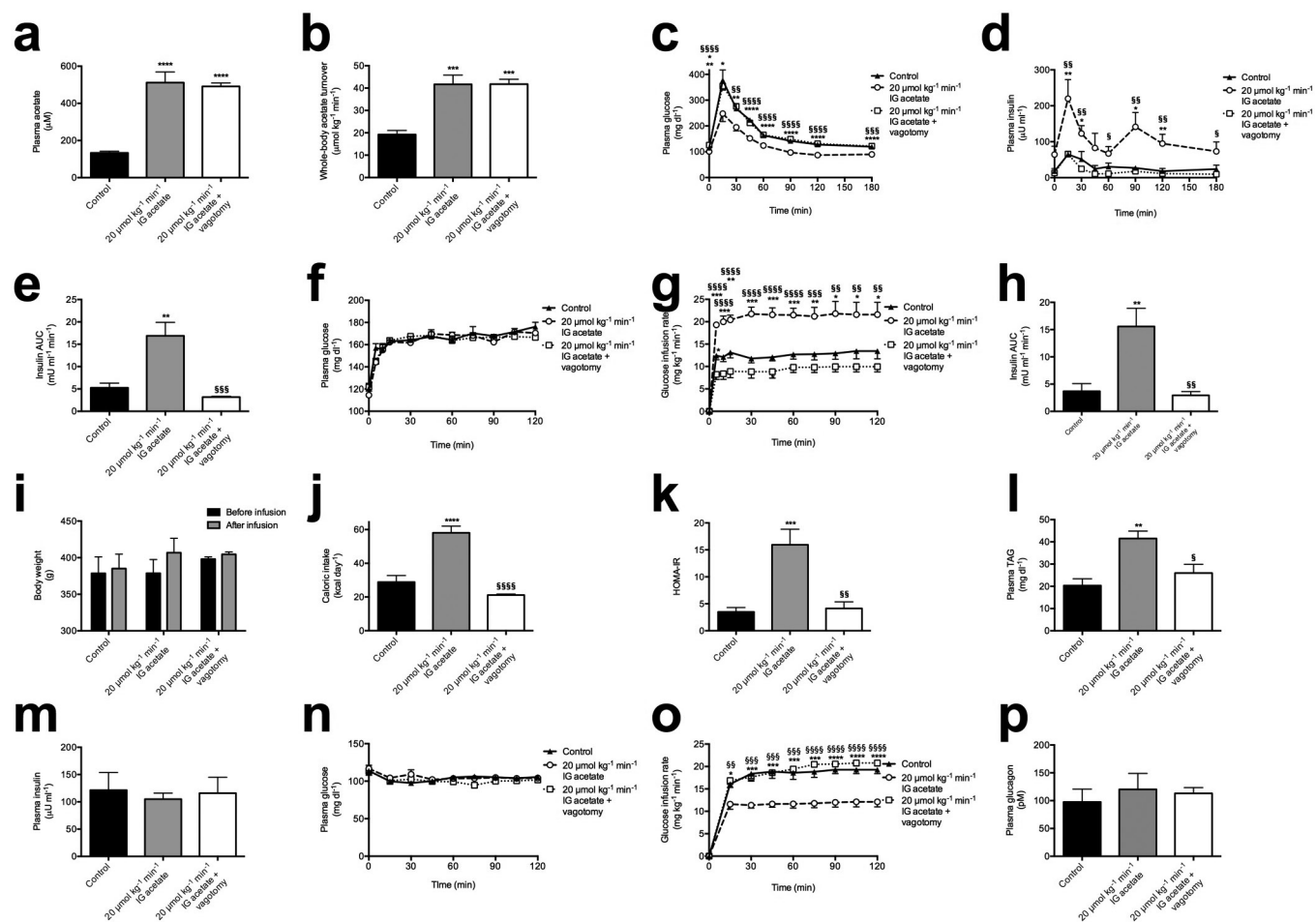
clamp in acetate-infused, atropine-treated rats. **k**, **l**, Plasma acetate and whole-body acetate turnover. **m**, Plasma insulin AUC during the clamp. **n**, Plasma glucagon. In all panels, \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$  by the two-tailed unpaired Student's *t*-test; data represent the mean  $\pm$  s.e.m. of  $n = 6$  replicates per group.



**Extended Data Figure 8 | Acetate drives GSIS via parasympathetic activation.** **a, b**, Plasma glucose and glucose infusion rate during a hyperglycaemic clamp in rats treated with ICV acetate. **c**, Plasma insulin AUC. **d**, Plasma glucagon. **e, f**, Plasma acetate and whole-body acetate turnover in rats treated with systemic intra-arterial acetate and ICV methylatropine. **g, h**, Plasma glucose and glucose infusion rate during a hyperglycaemic clamp. **i**, Plasma insulin AUC during the clamp. **j, k**, Plasma and brain tissue acetate in rats given an injection of acetate into the nucleus tractus solitarius. **l, m**, Plasma glucose and glucose

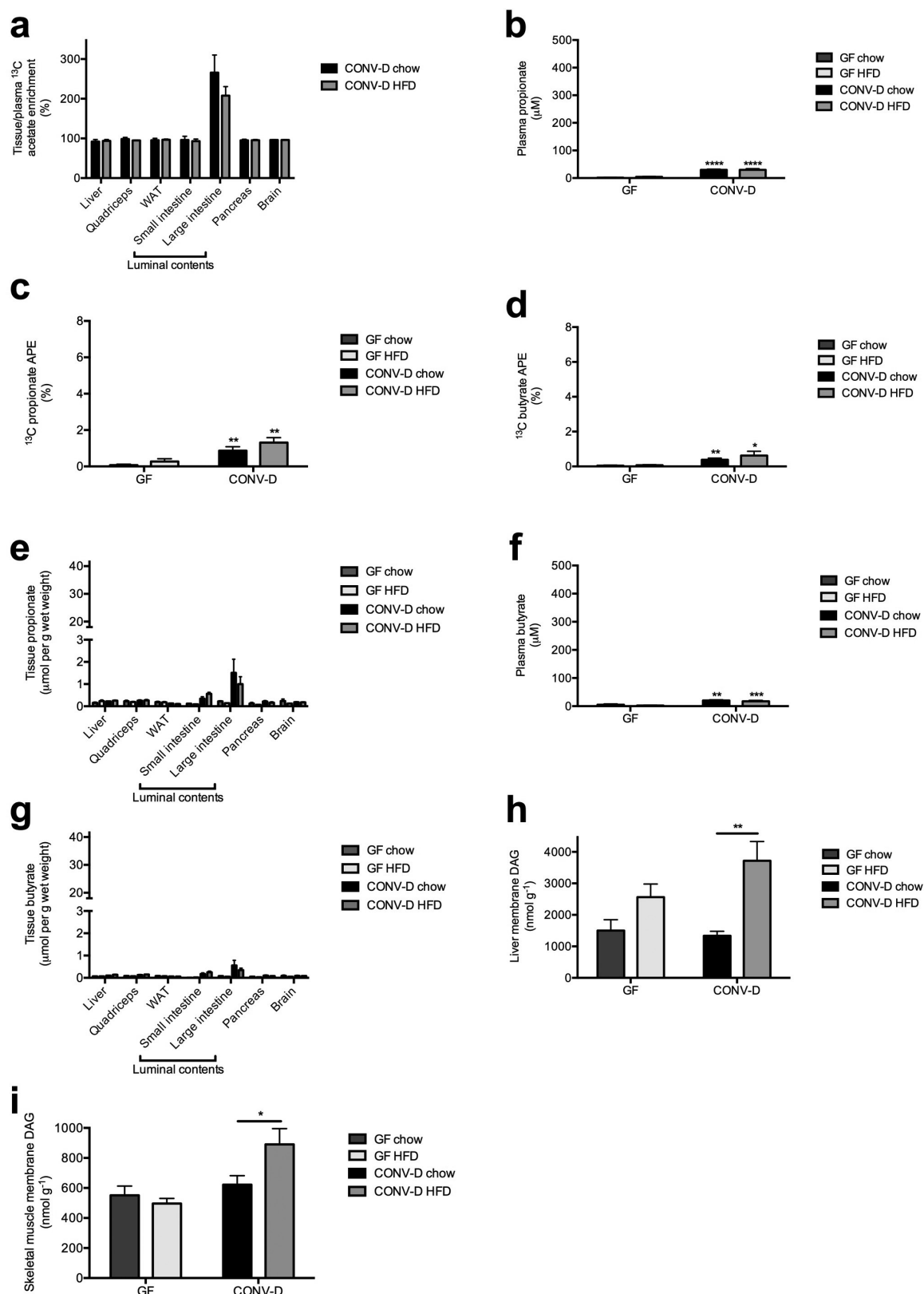
infusion rate during a hyperglycaemic clamp. **n**, Plasma insulin AUC during the clamp. **o**, Plasma glucagon. In all panels, data are the mean  $\pm$  s.e.m. of  $n = 6$  animals per group, with comparisons by one-way ANOVA with Bonferroni's multiple comparisons test (**a–i**) or two-tailed unpaired Student's *t*-test (**j–o**). In **b–d**, \*\* $P < 0.01$ , \*\*\* $P < 0.001$  versus controls; §§ $P < 0.01$ , §§§ $P < 0.001$  versus ICV acetate-treated rats by one-way ANOVA with Bonferroni's multiple comparisons test. In **e–i**, \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$  versus controls; §§§ $P < 0.001$ , §§§§ $P < 0.0001$  versus acetate-infused rats.





**Extended Data Figure 9 | Chronic intra-gastric acetate infusion causes hyperphagia and metabolic syndrome through parasympathetic activation.** **a, b**, Plasma acetate and whole-body acetate turnover. **c, d**, Plasma glucose and insulin concentrations during an intraperitoneal glucose tolerance test. **e**, Insulin AUC during the glucose tolerance test. **f, g**, Plasma glucose and glucose infusion rate during a hyperglycaemic clamp. **h**, Insulin AUC during the hyperglycaemic clamp. **i**, Body weight before and after the infusion study ( $n = 16$  controls, 16 acetate-infused, and 12 acetate-infused and vagotomised rats). **j**, Caloric intake during the 10-day acetate infusion study. **k**, Homeostatic model assessment of

insulin resistance (HOMA-IR). **l**, Plasma triglyceride concentrations. **m**, Plasma insulin at the 120-min time point of a hyperinsulinaemic-euglycaemic clamp. **n, o**, Plasma glucose and glucose infusion rate during the hyperinsulinaemic-euglycaemic clamp. **p**, Plasma glucagon. Unless otherwise specified, data are mean  $\pm$  s.e.m. of  $n = 6$  rats per group, with comparisons by one-way ANOVA with Bonferroni's multiple comparisons test. In all panels,  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ ,  $****P < 0.0001$  versus controls;  $\$P < 0.05$ ,  $\$\$P < 0.01$ ,  $\$\$\$P < 0.001$ ,  $\$\$\$\$P < 0.0001$  versus intra-gastric acetate-infused rats.



**Extended Data Figure 10 | Germ-free mice have negligible endogenous short-chain fatty acid production.** **a**, Ratio of tissue:plasma  $^{13}\text{C}$  acetate in mice fed  $^{13}\text{C}$  bicarbonate. **b**, **c**, Plasma and tissue propionate concentrations. **d**, Plasma  $^{13}\text{C}$  propionate enrichment. **e**, **f**, Plasma and tissue butyrate. **g**, Plasma  $^{13}\text{C}$  butyrate enrichment. **h**, **i**, Liver and muscle

diacylglycerol concentrations. In all panels, data are the mean  $\pm$  s.e.m. of  $n = 9$  (GF) or  $n = 10$  (CONV-D) mice per group, with comparisons by two-tailed unpaired Student's *t*-test. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$  versus CONV-D mice on the same diet.

# Cold, clumpy accretion onto an active supermassive black hole

Grant R. Tremblay<sup>1,2</sup>, J. B. Raymond Oonk<sup>3,4</sup>, Françoise Combes<sup>5</sup>, Philippe Salomé<sup>5</sup>, Christopher P. O'Dea<sup>6,7</sup>, Stefi A. Baum<sup>6,8</sup>, G. Mark Voit<sup>9</sup>, Megan Donahue<sup>9</sup>, Brian R. McNamara<sup>10</sup>, Timothy A. Davis<sup>2,11</sup>, Michael A. McDonald<sup>12</sup>, Alastair C. Edge<sup>13</sup>, Tracy E. Clarke<sup>14</sup>, Roberto Galván-Madrid<sup>2,15</sup>, Malcolm N. Bremer<sup>16</sup>, Louise O. V. Edwards<sup>1</sup>, Andrew C. Fabian<sup>17</sup>, Stephen Hamer<sup>5</sup>, Yuan Li<sup>18</sup>, Anaëlle Maury<sup>19</sup>, Helen R. Russell<sup>17</sup>, Alice C. Quillen<sup>20</sup>, C. Megan Urry<sup>1</sup>, Jeremy S. Sanders<sup>21</sup> & Michael W. Wise<sup>3</sup>

**Supermassive black holes in galaxy centres can grow by the accretion of gas, liberating energy that might regulate star formation on galaxy-wide scales<sup>1–3</sup>. The nature of the gaseous fuel reservoirs that power black hole growth is nevertheless largely unconstrained by observations, and is instead routinely simplified as a smooth, spherical inflow of very hot gas<sup>4</sup>. Recent theory<sup>5–7</sup> and simulations<sup>8–10</sup> instead predict that accretion can be dominated by a stochastic, clumpy distribution of very cold molecular clouds—a departure from the ‘hot mode’ accretion model—although unambiguous observational support for this prediction remains elusive. Here we report observations that reveal a cold, clumpy accretion flow towards a supermassive black hole fuel reservoir in the nucleus of the Abell 2597 Brightest Cluster Galaxy (BCG), a nearby (redshift  $z=0.0821$ ) giant elliptical galaxy surrounded by a dense halo of hot plasma<sup>11–13</sup>. Under the right conditions, thermal instabilities produce a rain of cold clouds that fall towards the galaxy’s centre<sup>14</sup>, sustaining star formation amid a kiloparsec-scale molecular nebula that is found at its core<sup>15</sup>. The observations show that these cold clouds also fuel black hole accretion, revealing ‘shadows’ cast by the molecular clouds as they move inward at about 300 kilometres per second towards the active supermassive black hole, which serves as a bright backlight. Corroborating evidence from prior observations<sup>16</sup> of warmer atomic gas at extremely high spatial resolution<sup>17</sup>, along with simple arguments based on geometry and probability, indicate that these clouds are within the innermost hundred parsecs of the black hole, and falling closer towards it.**

We observed the Abell 2597 BCG (Fig. 1) with the Atacama Large Millimeter/submillimeter Array (ALMA), enabling us to create a three-dimensional map of both the location and motions of cold gas at uniquely high sensitivity and spatial resolution. The ALMA receivers were sensitive to emission from the  $J=2-1$  rotational line of the carbon monoxide (CO) molecule. Such CO(2–1) emission is used as a tracer of cold ( $\sim 10$ – $30$  K) molecular hydrogen, which is vastly more abundant, but not directly observable at these low temperatures.

The continuum-subtracted CO(2–1) images (Fig. 2) reveal that the filamentary emission line nebula that spans the galaxy’s innermost  $\sim 30$  kpc (Fig. 1b) consists not only of warm ionized gas<sup>18–20</sup>, but also cold molecular gas. In projection, the optical emission line nebula is

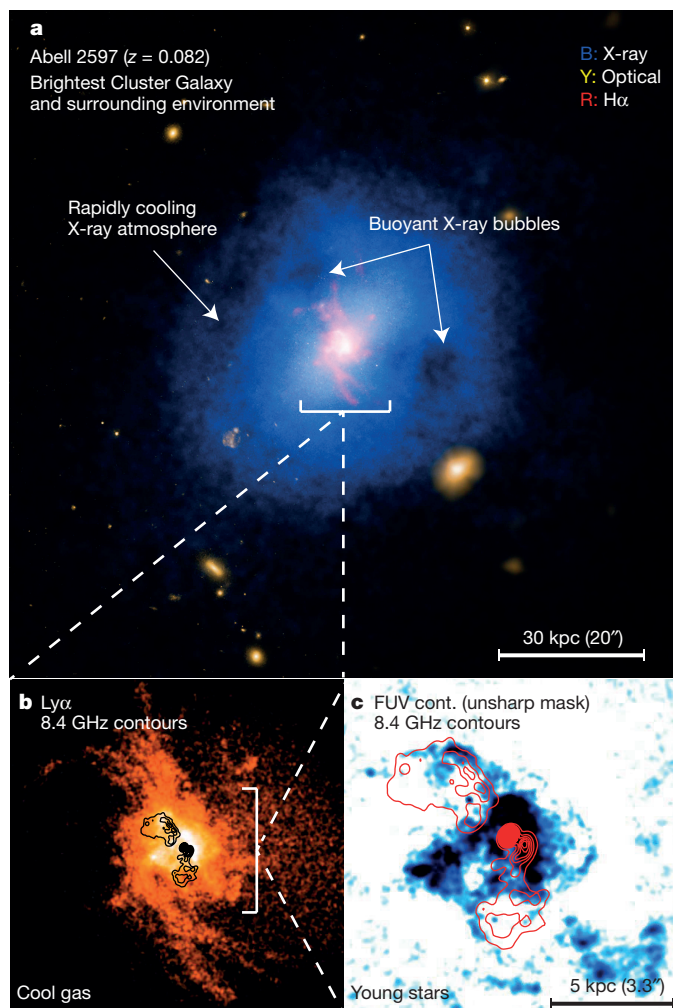
co-spatial and morphologically matched with CO(2–1) emission detected at a significance between  $\gtrsim 3\sigma$  (in the outer filaments) and  $\gtrsim 20\sigma$  (in the nuclear region) above the background noise level. The warm ionized nebula is therefore likely to have a substantial molecular component, consistent with results for other similar galaxies<sup>21</sup>. The total measured CO(2–1) line flux corresponds to a molecular hydrogen gas mass of  $M_{\text{H}_2} = (1.8 \pm 0.2) \times 10^9 M_\odot$ , where  $M_\odot$  is the mass of the Sun. The critical (minimum) density for CO(2–1) emission requires that the volume filling factor of this gas be very low, of the order of a few per cent. The projected spatial coincidence of both the warm ionized and cold molecular nebulae therefore supports the long-envisaged hypothesis that the ionized gas is merely the warm ‘skin’ surrounding far colder and more massive molecular cores<sup>22,23</sup>, whose outer regions are heated by intense radiation from the environment in which they reside. Rather than a monolithic, kiloparsec-scale slab of cold gas, we are more likely to be observing a projected superposition of many smaller, isolated clouds and filaments.

The data unambiguously show that cold molecular gas is falling inward along a line of sight that intersects the galaxy centre. We know this because the ALMA beam that is co-spatial with the millimetre continuum source, the radio core, and the isophotal centre of the galaxy reveals strong, redshifted continuum absorption (Fig. 3b), found by extracting the CO(2–1) spectrum from this central beam. This reveals at least three deep and narrow absorption lines (Fig. 3c), with redshifted line centres at  $+240$ ,  $+275$ , and  $+335 \text{ km s}^{-1}$  relative to the systemic (stellar) velocity of the galaxy, all within an angular (physical) region of  $0.715'' \times 0.533''$  ( $1 \text{ kpc} \times 0.8 \text{ kpc}$ ).

These absorption features arise from cold molecular clouds moving towards the centre of the galaxy, via either radial or inspiralling trajectories. They manifest as continuum absorption because they cast ‘shadows’ along the line of sight as the clouds eclipse or attenuate about  $\sim 20\%$  (or about 2 mJy) of the millimetre synchrotron continuum source, which serves as a bright backlight (13.6 mJy at rest-frame 230 GHz). The synchrotron continuum is emitted by jets launched from the accreting supermassive ( $\sim 3 \times 10^8 M_\odot$ ; ref. 13) black hole in the galaxy’s active nucleus (Fig. 4). The absorbers must therefore be located somewhere between the observer and the galaxy centre, falling deeper into the galaxy at about  $+300 \text{ km s}^{-1}$  towards the black hole at its core.

<sup>1</sup>Yale Center for Astronomy and Astrophysics, Yale University, 52 Hillhouse Avenue, New Haven, Connecticut 06511, USA. <sup>2</sup>European Southern Observatory, Karl-Schwarzschild-Strasse 2, 85748 Garching bei München, Germany. <sup>3</sup>ASTRON, Netherlands Institute for Radio Astronomy, PO Box 2, 7990 AA Dwingeloo, The Netherlands. <sup>4</sup>Leiden Observatory, Leiden University, Niels Borhweg 2, NL-2333 CA Leiden, The Netherlands. <sup>5</sup>Laboratory for Studies of Radiation and Matter in Astrophysics and Atmospheres (LERMA), Observatoire de Paris, PSL Research University, Collège de France, CNRS, Sorbonne University, Paris, France. <sup>6</sup>Department of Physics and Astronomy, University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada. <sup>7</sup>School of Physics and Astronomy, Rochester Institute of Technology, 84 Lomb Memorial Drive, Rochester, New York 14623, USA. <sup>8</sup>Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, 84 Lomb Memorial Drive, Rochester, New York 14623, USA. <sup>9</sup>Physics and Astronomy Department, Michigan State University, East Lansing, Michigan 48824-2320, USA. <sup>10</sup>Physics and Astronomy Department, Waterloo University, 200 University Avenue West, Waterloo, Ontario N2L 2G1, Canada. <sup>11</sup>School of Physics and Astronomy, Cardiff University, The Parade, Cardiff CF24 3AA, UK. <sup>12</sup>Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA. <sup>13</sup>Department of Physics, Durham University, Durham DH1 3LE, UK. <sup>14</sup>Naval Research Laboratory Remote Sensing Division, Code 7213, 4555 Overlook Avenue Southwest, Washington DC 20375, USA. <sup>15</sup>Instituto de Radioastronomía y Astrofísica, UNAM, Apartado Postal 3-72 (Xangari), 58089 Morelia, Michoacán, Mexico. <sup>16</sup>H. W. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK. <sup>17</sup>Institute of Astronomy, Cambridge University, Madingley Road, Cambridge CB3 0HA, UK. <sup>18</sup>Department of Astronomy, University of Michigan, 1085 South University Avenue, Ann Arbor, Michigan 48109, USA. <sup>19</sup>Laboratoire AIM-Paris-Saclay, CEA/DSM/Irfu CNRS—Université Paris Diderot, CE-Saclay, F-91191 Gif-sur-Yvette, France. <sup>20</sup>Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627, USA. <sup>21</sup>Max-Planck-Institut für Extraterrestrische Physik, 85748 Garching bei München, Germany.

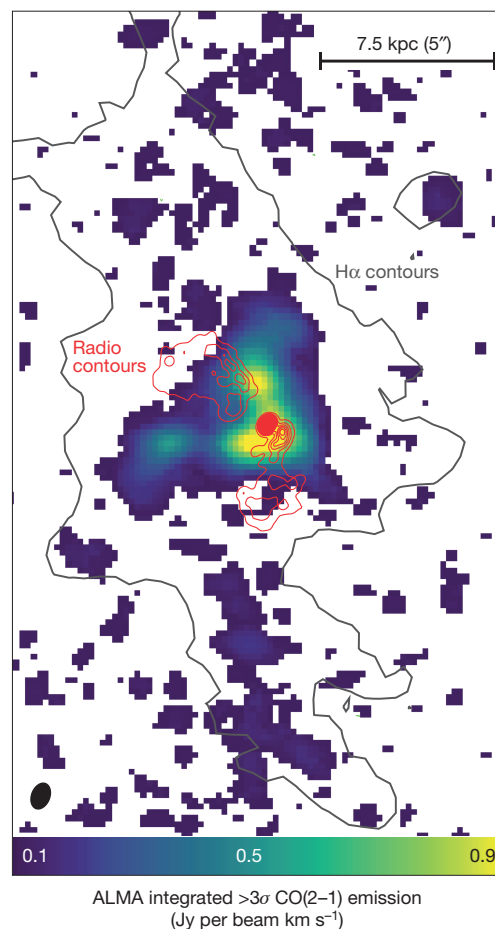




**Figure 1 | A multiwavelength view of the Abell 2597 BCG.** **a**, Chandra X-ray, HST and DSS optical, and Magellan  $H\alpha + [N II]$  emission is shown in blue, yellow and red, respectively. Arrows point to the thermally unstable hot atmosphere and buoyant bubbles that permeate it. (Image credits: X-ray, NASA/CXC/Michigan State University/G. Voit *et al.*; optical, NASA/STScI and DSS;  $H\alpha$ , Carnegie Observatory/Magellan/W. Baade Telescope/University of Maryland/M. McDonald). **b**, HST image of  $Ly\alpha$  emission associated with the ionized gas nebula<sup>13</sup>, with 8.4 GHz radio contours overlaid in black. The filamentary ionized nebula consists of cooler gas that has precipitated from the hot X-ray bright halo shown in **a**. **c**, Unsharp mask of the HST far-ultraviolet (FUV) continuum image of the central regions of the nebula<sup>15</sup>. The FUV emission directly traces the locations of young stars in the nebula. Very Large Array (VLA) radio contours of the 8.4 GHz source are overlaid in red.

This radial speed is roughly equal to the expected circular velocity<sup>24</sup> in the nucleus, consistent either with a nearly radial orbit, or with highly non-circular motions in close proximity to the galaxy's core.

Gaussian fits to the spectral absorption features reveal narrow line-widths of  $\sigma_v \lesssim 6 \text{ km s}^{-1}$ , which means the absorbers are more probably spatially compact, with sizes that span tens (rather than hundreds or thousands) of parsecs. The shapes of the absorption lines remain roughly the same regardless of how finely the spectra are binned, suggesting that the absorbers are probably coherent structures, rather than a superposition of many smaller absorbers unresolved in velocity space. If each absorption feature corresponds to one coherent cloud, and if those clouds roughly obey size–linewidth relations<sup>25,26</sup> for giant molecular clouds in the Milky Way, they should have diameters not larger than  $\sim 40 \text{ pc}$ . If in virial equilibrium, molecular clouds this size would have masses of the order of  $10^5$ – $10^6 M_\odot$ , and if in rough pressure equilibrium with their ambient multiphase ( $10^3$ – $10^7 \text{ K}$ ) environment<sup>13</sup>, they

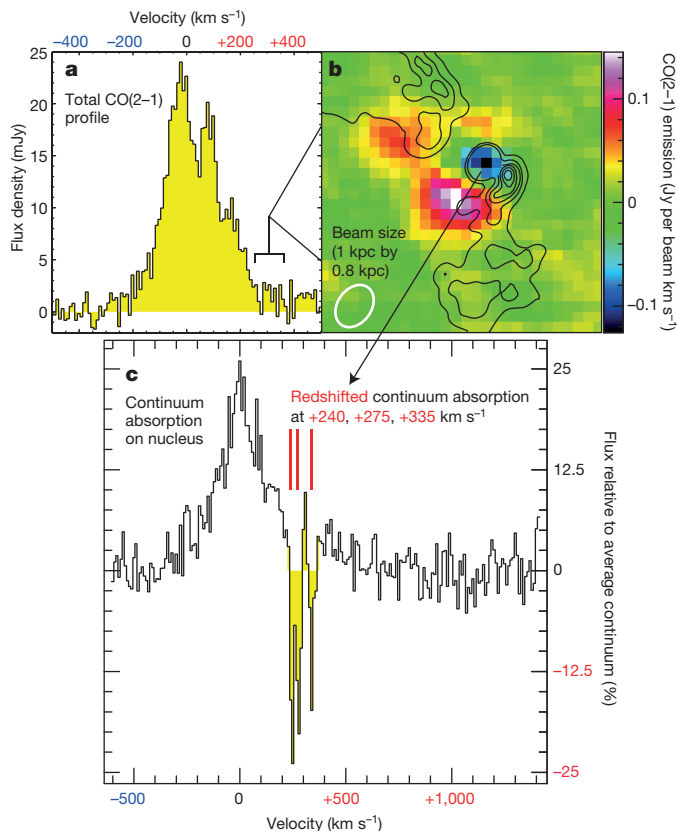


**Figure 2 | ALMA observation of continuum-subtracted CO(2–1) emission in the Abell 2597 BCG.** Emission is integrated from  $-600 \text{ km s}^{-1}$  to  $+600 \text{ km s}^{-1}$  relative to the galaxy's systemic velocity. Channels are binned to  $40 \text{ km s}^{-1}$ . Only  $\geq 3\sigma$  emission is shown. 8.4 GHz VLA radio contours are overlaid in black, and  $H\alpha$  contours outlining the rough boundary of the ionized nebula are shown in grey. The nebula is slightly larger than the grey contours suggest: emission outside this boundary is still part of a smooth, fainter distribution of cold gas, co-spatial with similarly faint emission in the optical.

must have high column densities of the order of  $N_{H_2} \approx 10^{22}$ – $10^{24} \text{ cm}^{-2}$  so as to maintain pressure support. The thermal pressure in the core of Abell 2597 BCG is nearly 3,000 times<sup>13</sup> greater than that for the Milky Way, however, which means the absorbing clouds may be much smaller.

The absorbers have optical depths in the range  $0.1 \lesssim \tau_{CO(2-1)} \lesssim 0.3$ . The physical resolution of the ALMA data is larger than the synchrotron background source, which means that the optical depth is probably contaminated by an unresolved, additive superposition of both emission and absorption within the beam. Compact, dense cold clouds are nevertheless likely to be optically thick, which may mean they eclipse the continuum source with an optical depth of unity but a small covering factor of roughly 0.2. Especially when considering beam contamination by emission, the covering factor cannot be known with certainty, as this depends on the unknown geometry of the absorbing and emitting regions within the ALMA beam.

This geometry can be constrained, however, given existing Very Long Baseline Array (VLBA) radio observations at extremely high spatial resolution<sup>17</sup>. These data resolve the 1.3 GHz and 5 GHz radio continuum source down to scales of 25 pc, revealing a highly symmetric, 100-pc-scale jet about a bright radio core (Fig. 4c). Just as we have found in cold molecular gas, inflowing warmer atomic hydrogen gas ( $H I$ ) has previously been found in absorption against this parsec-scale jet, corroborating prior reports of inflowing atomic gas at lower spatial resolutions<sup>16</sup>. The inflow velocity of this gas matches that seen

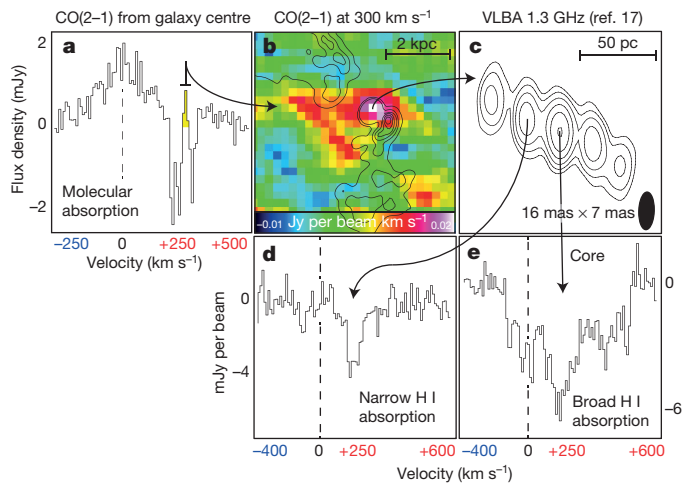


**Figure 3 | ‘Shadows’ cast by molecular clouds moving towards the supermassive black hole.** **a**, Continuum-subtracted ALMA CO(2-1) spectrum extracted from a central 10 kpc region. Brackets mark CO(2-1) emission shown in **b**, where 8.4 GHz radio contours are overlaid. The central radio contours have been removed to aid viewing of the continuum absorption, seen as the blue/black spot of ‘negative’ emission (which is the radio and mm core, the centre of the galaxy and the location of the black hole). **c**, Continuum-subtracted CO(2-1) spectrum extracted from this region co-spatial with the mm and radio core. Absorption lines are indicated in red.

in our ALMA data. Remarkably, both the optical depth and linewidth of the warm atomic absorption signal varies dramatically across the jet, with a broad ( $\sigma_v \approx 310 \text{ km s}^{-1}$ ) component co-spatial with the core that is absent just  $\sim 20 \text{ pc}$  to the northeast, where only a narrow ( $\sigma_v \approx 50 \text{ km s}^{-1}$ ) H I line is found at the same redshift. This effectively requires the inflowing atomic gas to be confined within the innermost  $\sim 100 \text{ pc}$  of the black hole, as gas further out would give rise to an unchanging absorption signal across the compact jet. The infall velocity is the same as that for the cold molecular clouds seen in CO(2-1) absorption, which means they most probably stem from the same spatial region, within tens of parsecs of the accreting black hole.

This is further supported by the ALMA data itself. In emission, all gas around approximately  $+300 \text{ km s}^{-1}$  that is conceivably available to attenuate the continuum signal is confined to the innermost 2 kpc about the nucleus (Fig. 4a, b). The radial dependence of molecular cloud volume number density within this region is uncertain, but probably steeper than  $r^{-1}$ , and likely to be closer to  $r^{-2}$  (Fig. 4b). This means that the chances of a random line of sight crossing will drop with increasing distance from the black hole. If the gas volume density goes as  $r^{-2}$ , a cloud 100 pc from the black hole is ten times more likely to cross our line of sight than a cloud at a galactocentric distance of 1 kpc. It would be exceedingly unlikely for three such clouds to cross our line of sight to the black hole were they spread over several kiloparsecs throughout the galaxy’s outskirts.

The data therefore serve as strong observational evidence for an inward-moving, clumpy distribution of molecular clouds within a few



**Figure 4 | Corroborating evidence that the inflowing molecular clouds must be in close proximity to the black hole.** **a**, CO(2-1) absorption spectrum from Fig. 3, with a region of emission at about  $+300 \text{ km s}^{-1}$  marked in yellow. **b**, Integrated CO(2-1) emission (colour coded) from this region, showing that gas at about  $+300 \text{ km s}^{-1}$  is confined to the innermost 2 kpc of the galaxy. **c**, 1.3 GHz radio continuum source from an archival VLBA observation<sup>17</sup> with an extremely high physical resolution of  $\sim 25 \text{ pc}$  by  $\sim 10 \text{ pc}$ . **d**, **e**, Plots of 1.3 GHz radio continuum emission revealing H I 21 cm absorption observed against this synchrotron jet. The signal varies dramatically over scales of tens of parsecs.

hundred parsecs of an accreting supermassive black hole. The result augments a small but growing set of known molecular absorption systems<sup>27–29</sup> whose black hole proximity is less well constrained. The infalling clouds in Abell 2597 BCG are probably a few to tens of parsecs across and therefore massive (perhaps  $10^5$ – $10^6 M_\odot$  each). If they are falling directly towards the black hole, rather than bound in a non-circular orbit that tightly winds around it, they could supply an upper-limit accretion rate of the order of  $\sim 0.1$  to a few solar masses per year, depending on the three-dimensional distribution of infalling clouds. If most of the clouds are instead locked in non-circular orbits around the black hole, the fuelling rate would depend on the gas angular momentum, and the local supply of torques that might lessen it. Simulations suggest<sup>9,10,14</sup> that such torques may be plentiful, as they predict a stochastic ‘rain’ of thermal instabilities that condense from all directions around the black hole, promoting angular momentum cancellation via tidal stress and cloud–cloud collisions. Even highly elliptical cloud orbits should therefore be associated with significant inward radial motions. The clouds might fall onto the accretion disk itself, or into a clumpy rotating ring akin to the ‘torus’ invoked in AGN unification models<sup>30</sup>.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 17 December 2015; accepted 22 March 2016.**

- McNamara, B. R. & Nulsen, P. E. J. Heating hot atmospheres with active galactic nuclei. *Annu. Rev. Astron. Astrophys.* **45**, 117–175 (2007).
- McNamara, B. R. & Nulsen, P. E. J. Mechanical feedback from active galactic nuclei in galaxies, groups and clusters. *New J. Phys.* **14**, 055023 (2012).
- Fabian, A. C. Observational evidence of active galactic nuclei feedback. *Annu. Rev. Astron. Astrophys.* **50**, 455–489 (2012).
- Bondi, H. On spherically symmetrical accretion. *Mon. Not. R. Astron. Soc.* **112**, 195–204 (1952).
- Pizzolato, F. & Soker, N. On the nature of feedback heating in cooling flow clusters. *Astrophys. J.* **632**, 821–830 (2005).
- Voit, G. M., Donahue, M., Bryan, G. L. & McDonald, M. Regulation of star formation in giant galaxies by precipitation, feedback and conduction. *Nature* **519**, 203–206 (2015).
- Voit, G. M., Bryan, G. L., O’Shea, B. W. & Donahue, M. Precipitation-regulated star formation in galaxies. *Astrophys. J.* **808**, L30 (2015).

8. Sharma, P., McCourt, M., Quataert, E. & Parrish, I. J. Thermal instability and the feedback regulation of hot haloes in clusters, groups, and galaxies. *Mon. Not. R. Astron. Soc.* **420**, 3174–3194 (2012).
9. Gaspari, M., Ruszkowski, M. & Oh, S. P. Chaotic cold accretion on to black holes. *Mon. Not. R. Astron. Soc.* **432**, 3401–3422 (2013).
10. Li, Y. & Bryan, G. L. Modeling active galactic nucleus feedback in cool-core clusters: the formation of cold clumps. *Astrophys. J.* **789**, 153 (2014).
11. McNamara, B. R. *et al.* Discovery of ghost cavities in the X-ray atmosphere of Abell 2597. *Astrophys. J.* **562**, L149–L152 (2001).
12. Clarke, T. E., Sarazin, C. L., Blanton, E. L., Neumann, D. M. & Kassim, N. E. Low-frequency radio observations of X-ray ghost bubbles in A2597: a history of radio activity in the core. *Astrophys. J.* **625**, 748–753 (2005).
13. Tremblay, G. R. *et al.* Multiphase signatures of active galactic nucleus feedback in Abell 2597. *Mon. Not. R. Astron. Soc.* **424**, 1026–1041 (2012).
14. Gaspari, M., Brighenti, F. & Temi, P. Chaotic cold accretion on to black holes in rotating atmospheres. *Astron. Astrophys.* **579**, A62 (2015).
15. Tremblay, G. R. *et al.* Residual cooling and persistent star formation amid active galactic nucleus feedback in Abell 2597. *Mon. Not. R. Astron. Soc.* **424**, 1042–1060 (2012).
16. O'Dea, C. P., Baum, S. A. & Gallimore, J. F. Detection of extended H I absorption toward PKS 2322–123 in Abell 2597. *Astrophys. J.* **436**, 669–677 (1994).
17. Taylor, G. B., O'Dea, C. P., Peck, A. B. & Koekemoer, A. M. H I Absorption toward the nucleus of the radio galaxy PKS 2322–123 in A2597. *Astrophys. J.* **512**, L27–L30 (1999).
18. O'Dea, C. P., Baum, S. A., Mack, J., Koekemoer, A. M. & Laor, A. Hubble Space Telescope STIS far-ultraviolet observations of the central nebulae in the cooling-core clusters A1795 and A2597. *Astrophys. J.* **612**, 131–151 (2004).
19. Oonk, J. B. R., Hatch, N. A., Jaffe, W., Bremer, M. N. & van Weeren, R. J. Far-ultraviolet emission in the A2597 and A2204 brightest cluster galaxies. *Mon. Not. R. Astron. Soc.* **414**, 2309–2336 (2011).
20. Tremblay, G. R. *et al.* Far-ultraviolet morphology of star-forming filaments in cool core brightest cluster galaxies. *Mon. Not. R. Astron. Soc.* **451**, 3768–3800 (2015).
21. Salomé, P. *et al.* A very extended molecular web around NGC 1275. *Astron. Astrophys.* **531**, A85 (2011).
22. Jaffe, W., Bremer, M. N. & Baker, K. H II and H<sub>2</sub> in the envelopes of cooling flow central galaxies. *Mon. Not. R. Astron. Soc.* **360**, 748–762 (2005).
23. Salomé, P. *et al.* Cold molecular gas in the Perseus cluster core. Association with X-ray cavity, H $\alpha$  filaments and cooling flow. *Astron. Astrophys.* **454**, 437–445 (2006).
24. Smith, E. P., Heckman, T. M. & Illingworth, G. D. Stellar dynamics of powerful radio galaxies. *Astrophys. J.* **356**, 399–415 (1990).
25. Larson, R. B. Turbulence and star formation in molecular clouds. *Mon. Not. R. Astron. Soc.* **194**, 809–826 (1981).
26. Solomon, P. M., Rivolo, A. R., Barrett, J. & Yahil, A. Mass, luminosity, and line width relations of Galactic molecular clouds. *Astrophys. J.* **319**, 730–741 (1987).
27. Wiklind, T. & Combes, F. Molecular absorption and its time variations in Centaurus A. *Astron. Astrophys.* **324**, 51–64 (1997).
28. Espada, D. *et al.* Disentangling the circumnuclear environs of Centaurus A. II. On the nature of the broad absorption line. *Astrophys. J.* **720**, 666–678 (2010).
29. David, L. P. *et al.* Molecular gas in the X-ray bright group NGC 5044 as revealed by ALMA. *Astrophys. J.* **792**, 94 (2014).
30. Urry, C. M. & Padovani, P. Unified schemes for radio-loud active galactic nuclei. *Publ. Astron. Soc. Pacif.* **107**, 803–845 (1995).

**Acknowledgements** ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada) and NSC and ASIAA (Taiwan), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ. We are grateful to the European ALMA Regional Centres, particularly those in Garching and Manchester, for their dedicated end-to-end support of data associated with this paper. We thank R. Larson for discussions. G.R.T. acknowledges support from National Aeronautics and Space Administration (NASA) through Einstein Postdoctoral Fellowship Award Number PF-150128, issued by the Chandra X-ray Observatory Center, which is operated by the Smithsonian Astrophysical Observatory for and on behalf of NASA under contract NAS8-03060. F.C. acknowledges the European Research Council (ERC) for the Advanced Grant Program no. 267399-Momentum. B.R.M. is supported by a grant from the Natural Sciences and Engineering Research Council of Canada. T.A.D. acknowledges support from a Science and Technology Facilities Council (STFC) Ernest Rutherford Fellowship. A.C.E. acknowledges support from STFC grant ST/L00075X/1. A.C.F. and H.R.R. acknowledge support from ERC Advanced Grant Program no. 340442-Feedback. M.N.B. acknowledges funding from the STFC. Basic research in radio astronomy at the Naval Research Laboratory is supported by 6.1 Base funding.

**Author Contributions** G.R.T. was principal investigator on the original proposal, performed the data analysis, and wrote the paper. J.B.R.O., T.A.D., R.G.M. and A.M. were substantially involved in planning both scientific and technical aspects of the proposal, while T.A.D. and R.G.M. contributed ALMA data reduction and analysis expertise once the data were obtained. J.B.R.O., F.C. and P.S. invested substantial time in analysis of the data. Substantial scientific feedback was also provided over many months by F.C., J.B.R.O., C.P.O., S.A.B., G.M.V., M.D., B.R.M., M.A.M., T.E.C., H.R., A.C.E. and A.C.F., while all other co-authors discussed the results and commented on the manuscript.

**Author Information** This paper makes use of the following ALMA data: ADS/JAO.ALMA#2012.1.00988.S. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.R.T. ([grant.tremblay@yale.edu](mailto:grant.tremblay@yale.edu)).



## METHODS

**Observations, data reduction, and analysis.** The new ALMA data presented in this paper were obtained in Cycle 1 with the use of 29 operational antennae in the 12 m Array. ALMA's Band 6 heterodyne receivers were tuned to a frequency of 213 GHz, sensitive to the  $J=2-1$  rotational line transition of carbon monoxide at the redshift of the Abell 2597 BCG ( $z=0.0821$ ). The ALMA correlator, set to Frequency Division Mode (FDM), delivered a bandwidth of 1,875 MHz (per baseband) with a 0.488 MHz channel spacing, for a maximum spectral resolution of about  $2 \text{ km s}^{-1}$ . One baseband was centred on the CO(2–1) emission line, while the other three sampled the local continuum. Maximum antenna baselines extended to  $\sim 1 \text{ km}$ , delivering an angular resolution at 213 GHz of  $\sim 0.7''$  within a  $\sim 28''$  primary beam (field of view). ALMA observed the Abell 2597 BCG, located at RA 23 h 25 min 20 s, dec.  $-12^\circ 07' 38''$  (J2000), for a total of  $\sim 3 \text{ h}$  over three separate scheduling blocks executed between 17 and 19 November 2013. The planet Neptune and quasars J2258–2758 and J2331–1556 were used for amplitude, flux, and phase calibration. The data were reduced using CASA version 4.2 with calibration and imaging scripts kindly provided by the ALMA Regional Centers (ARCs) in both Garching, Germany and Manchester, UK. Beyond the standard application of the phase calibrator solution, we iteratively performed self-calibration of the data using the galaxy's own continuum, yielding a  $\sim 14\%$  decrease in RMS noise to a final value of  $0.16 \text{ mJy per } 0.715'' \times 0.533''$  beam per  $40 \text{ km s}^{-1}$  channel. There is effectively no difference in CO(2–1) morphology between the self-calibrated and non-self-calibrated data cubes. Measurement sets were imaged using 'natural' visibility weighting and binning to either  $5 \text{ km s}^{-1}$ ,  $10 \text{ km s}^{-1}$ , or  $40 \text{ km s}^{-1}$ , as indicated in the figure legends. The figures presented in this Letter show only continuum-subtracted, pure CO(2–1) line emission. The rest-frame 230 GHz continuum observation is dominated by a bright ( $13.6 \text{ mJy}$ ) point source associated with the AGN (detected at  $\gtrsim 400\sigma$ ), serving as the bright 'backlight' against which the continuum absorption features presented in this Letter were observed. The continuum data also features compact ( $\sim 5 \text{ kpc}$ ) extended emission at  $\sim 10\sigma$  that extends along the galaxy's dust lane, to be discussed in a forthcoming paper.

**Adoption of a systemic velocity.** Interpretation of gas motions relative to the stellar component of a galaxy requires adoption of a systemic (stellar) velocity to be used as a 'zero point' marking the transition from blue- to redshift. All CO(2–1) line velocities discussed in this Letter are set relative to  $213.04685 \text{ GHz}$ , where observed CO(2–1) emission peaks. This frequency corresponds to  $^{12}\text{CO}(2-1)$  (rest-frame  $230.538001 \text{ GHz}$ ) at a redshift of  $z=0.0821$ . This redshift is consistent, conservatively within  $\pm 60 \text{ km s}^{-1}$ , with every other available multiwavelength tracer of the galaxy's systemic velocity, including prominent Ca II H, K, and G-band absorption features<sup>17</sup> that directly trace the galaxy stellar component, the redshift of all optical emission lines<sup>31</sup>, as well as a broad (FWHM  $412 \text{ km s}^{-1}$ ) H I absorption component<sup>16</sup> at the optical emission and absorption line redshift. It is also consistent, within  $\sim 60 \text{ km s}^{-1}$ , with a cross-correlation of emission and absorption lines using galaxy template spectra<sup>17</sup>, as well as with all other published reports of the galaxy's systemic velocity (found, for example, within the HyperLeda database). We are therefore certain that the reported redshift of the absorption features discussed in this Letter indeed corresponds to real motion relative to the galaxy's stellar component. Without caveat or ambiguity, the absorbing cold clouds are moving into the galaxy at roughly  $\sim 300 \pm 60 \text{ km s}^{-1}$ .

**Mass estimates.** All molecular gas masses estimated in this letter adopt the following relation<sup>32</sup>:

$$M_{\text{mol}} = \frac{1.05 \times 10^4}{3.2} \left( \frac{X_{\text{CO}}}{2 \times 10^{20} \text{ cm}^{-2}} \right) \left( \frac{1}{1+z} \right) \left( \frac{S_{\text{CO}} \Delta \nu}{\text{Jy km s}^{-1}} \right) \left( \frac{D_L}{\text{Mpc}} \right)^2 M_{\odot}$$

where  $S_{\text{CO}} \Delta \nu$  is the emission integral (effectively the total CO flux over the region of interest),  $z$  is the galaxy redshift ( $z=0.0821$ ), and  $D_L$  its luminosity distance ( $373.3 \text{ Mpc}$ ), for which we assume a flat  $\Lambda$ CDM model wherein  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_M = 0.3$ , and  $\Omega_\Lambda = 0.7$ . This mass estimate most critically relies on an assumption of the CO-to- $\text{H}_2$  conversion factor<sup>32</sup>,  $X_{\text{CO}}$ . In this Letter we assume the average Milky Way value of  $X_{\text{CO}} = 2 \times 10^{20} \text{ cm}^{-2} (\text{K km s}^{-1})^{-1}$  and a CO(2–1) to CO(1–0) flux density ratio of 3.2. Other authors have provided extensive discussion of these assumptions as they pertain to cool core BCGs<sup>29,33,34</sup>. Scientific conclusions in this paper are largely insensitive to choice of  $X_{\text{CO}}$ .

A single Gaussian fit to the CO(2–1) spectrum extracted from an aperture containing all detected emission yields an emission integral of  $S_{\text{CO}} \Delta \nu = 4.2 \pm 0.4 \text{ Jy km s}^{-1}$  with a line FWHM of  $252 \pm 14 \text{ km s}^{-1}$ , corresponding to a total molecular hydrogen ( $\text{H}_2$ ) gas mass of  $M_{\text{H}_2} = (1.80 \pm 0.19) \times 10^9 M_{\odot}$ . This is very close to the previously reported<sup>15</sup> mass, based on an IRAM 30m CO(2–1) observation, of  $(1.8 \pm 0.3) \times 10^9 M_{\odot}$ . This comparison is not one-to-one, as the mass from the IRAM 30m observation was computed from within a beam size of  $11''$  (rather than  $28''$  for the ALMA data), and used a CO(2–1)/CO(1–0) flux ratio of 4 (rather than 3.2, as we use here). These differences are minor, particularly because nearly all of the CO(2–1) emission detected by ALMA is found within the central  $11''$  size of

the IRAM 30m beam. It is therefore safe to say that our ALMA observation has detected nearly all emission that was detected in the single-dish IRAM 30m observation, and that very little extended emission has been 'resolved out' by ALMA.

**Estimating physical properties of the redshifted absorbing molecular gas.** We have estimated a rough upper-limit size of the absorbing clouds assuming the widely adopted Larson *et al.*<sup>25</sup> and Solomon *et al.*<sup>26</sup> size–linewidth relation for molecular clouds in the Milky Way (namely, the ref. 26 fit);

$$\sigma_v = (1.0 \pm 0.1) S^{0.5 \pm 0.05} \text{ km s}^{-1}$$

where  $\sigma_v$  is the velocity linewidth of the cloud and  $S$  is the diameter of the cloud in parsecs. A measured absorber linewidth of  $\sigma_v \approx 6 \text{ km s}^{-1}$  would then correspond to a size of  $\sim 36 \text{ pc}$ . As noted in the main text of the Letter, the thermal pressure in the Abell 2597 BCG is about 3,000 times higher than that for the Milky Way<sup>13</sup>, so it is likely that the above relation does not apply. A higher ambient pressure implies higher compression and therefore smaller cloud size, so the above estimate should, at best, be considered a very rough upper-limit. The main lesson to take away from this exercise is that the absorbing clouds are probably physically compact (that is, a few to tens—rather than hundreds—of parsecs in diameter).

The three clouds are separated from one another by  $\sim 45\text{--}60 \text{ km s}^{-1}$  in velocity space, which means they are unlikely to be closely bound satellites of one another. Instead, it is more likely that they represent three random points along a radial distribution of clouds.

If the absorbers are in virial equilibrium, their masses  $M_{\text{cloud}}$  can be roughly estimated by applying the virial relation,

$$M_{\text{cloud}} \approx \frac{R_{\text{cloud}} \sigma_v^2}{G} \approx \frac{20 \text{ pc} \times (6 \text{ km s}^{-1})^2}{4.302 \times 10^{-3} \text{ pc } M_{\odot}^{-1} (\text{km s}^{-1})^2} \approx 1.7 \times 10^5 M_{\odot}$$

where  $R_{\text{cloud}}$  is the cloud radius (as roughly estimated above) and  $\sigma_v$  is its velocity dispersion (also as above).

CO(2–1) optical depths for the absorbers were estimated by assuming that:

$$I_{\text{total}} = I_{\text{continuum}} e^{-\tau_{\text{CO}(2-1)}}$$

where  $I_{\text{total}}$  and  $I_{\text{continuum}}$  are the integrated intensities of the total (line plus continuum) and continuum-only signals, respectively, and  $\tau_{\text{CO}(2-1)}$  is the optical depth of the CO(2–1) absorption feature.

The stellar velocity dispersion of the BCG<sup>33</sup> is  $\sigma_v = 220 \pm 19 \text{ km s}^{-1}$ . Under the assumption of an isothermal sphere, the circular velocity should be  $\sim 300 \text{ km s}^{-1}$  (that is,  $\sqrt{2} \sigma_v$ ), which is roughly the line of sight velocity of the absorption features. That the absorbers' redshift is a significant fraction of the expected circular velocity means they could be on a nearly radial orbit (though their transverse velocity cannot be known with this single observation).

While not discussed in the main text, there is an additional simple argument that independently suggests that the inward moving molecular clouds must be in close proximity to the black hole. If our line of sight is representative, and therefore a 'pencil beam' sample of a three-dimensional spherical distribution of clouds, the total mass of cold gas contained within this distribution should go roughly as:

$$M \approx 10^9 M_{\odot} \times f_c \times \left( \frac{r}{1 \text{ kpc}} \right)^2 \times \left( \frac{N_{\text{H}}}{10^{22} \text{ cm}^{-2}} \right)$$

where  $f_c$  is the covering factor and  $r$  is the radius of an imaginary thin spherical shell of molecular gas with column density  $N_{\text{H}}$ . If such a shell had a covering factor of 1, a radius of 1 kpc, and a column density of  $10^{22} \text{ cm}^{-2}$ , then the total mass of molecular hydrogen contained within that shell would be roughly one billion solar masses. A column density in excess of  $10^{22} \text{ cm}^{-2}$  requires this distribution to be contained within a sphere of radius  $< 1 \text{ kpc}$ , lest the limit set by the total mass of molecular hydrogen in the galaxy be violated. If the characteristic column density is  $10^{23} \text{ cm}^{-2}$ , for example, this mass must be contained within a sphere of radius 300 pc, or else its total mass would exceed the  $1.8 \times 10^9 M_{\odot}$  of cold gas present in the galaxy.

**Codes, software, and data availability.** Codes that we have written to both reduce and analyse the data presented in this Letter have been made publicly available at [https://github.com/grantremblay/Tremblay\\_Nature\\_ALMA\\_Abell2597](https://github.com/grantremblay/Tremblay_Nature_ALMA_Abell2597). Reduction of the data as well as some simple modelling (for example, fitting of Gaussians to lines) was performed using routines included in CASA version 4.2, available at <https://casa.nrao.edu/>. Plots were made using both Python's Matplotlib and *Veusz*, which is available at <http://home.gna.org/veusz/>.

31. Voit, G. M. & Donahue, M. A deep look at the emission-line nebula in Abell 2597. *Astrophys. J.* **486**, 242 (1997).
32. Bolatto, A. D., Wolfire, M. & Leroy, A. K. The CO-to- $\text{H}_2$  conversion factor. *Annu. Rev. Astron. Astrophys.* **51**, 207–268 (2013).
33. McNamara, B. R. *et al.* A  $10^{10}$  solar mass flow of molecular gas in the A1835 brightest cluster galaxy. *Astrophys. J.* **785**, 44 (2014).
34. Russell, H. R. *et al.* Massive molecular gas flows in the A1664 brightest cluster galaxy. *Astrophys. J.* **784**, 78 (2014).

# Digitized adiabatic quantum computing with a superconducting circuit

R. Barends<sup>1</sup>, A. Shabani<sup>2</sup>, L. Lamata<sup>3</sup>, J. Kelly<sup>1</sup>, A. Mezzacapo<sup>3,†</sup>, U. Las Heras<sup>3</sup>, R. Babbush<sup>2</sup>, A. G. Fowler<sup>1</sup>, B. Campbell<sup>4</sup>, Yu Chen<sup>1</sup>, Z. Chen<sup>4</sup>, B. Chiaro<sup>4</sup>, A. Dunsworth<sup>4</sup>, E. Jeffrey<sup>1</sup>, E. Lucero<sup>1</sup>, A. Megrant<sup>4</sup>, J. Y. Mutus<sup>1</sup>, M. Neeley<sup>1</sup>, C. Neill<sup>4</sup>, P. J. J. O'Malley<sup>4</sup>, C. Quintana<sup>4</sup>, P. Roushan<sup>1</sup>, D. Sank<sup>1</sup>, A. Vainsencher<sup>4</sup>, J. Wenner<sup>4</sup>, T. C. White<sup>4</sup>, E. Solano<sup>3,5</sup>, H. Neven<sup>2</sup> & John M. Martinis<sup>1,4</sup>

Quantum mechanics can help to solve complex problems in physics<sup>1</sup> and chemistry<sup>2</sup>, provided they can be programmed in a physical device. In adiabatic quantum computing<sup>3–5</sup>, a system is slowly evolved from the ground state of a simple initial Hamiltonian to a final Hamiltonian that encodes a computational problem. The appeal of this approach lies in the combination of simplicity and generality; in principle, any problem can be encoded. In practice, applications are restricted by limited connectivity, available interactions and noise. A complementary approach is digital quantum computing<sup>6</sup>, which enables the construction of arbitrary interactions and is compatible with error correction<sup>7,8</sup>, but uses quantum circuit algorithms that are problem-specific. Here we combine the advantages of both approaches by implementing digitized adiabatic quantum computing in a superconducting system. We tomographically probe the system during the digitized evolution and explore the scaling of errors with system size. We then let the full system find the solution to random instances of the one-dimensional Ising problem as well as problem Hamiltonians that involve more complex interactions. This digital quantum simulation<sup>9–12</sup> of the adiabatic algorithm consists of up to nine qubits and up to 1,000 quantum logic gates. The demonstration of digitized adiabatic quantum computing in the solid state opens a path to synthesizing long-range correlations and solving complex computational problems. When combined with fault-tolerance, our approach becomes a general-purpose algorithm that is scalable.

A key challenge in adiabatic quantum computing is to construct a device that is capable of encoding problem Hamiltonians that are classically intractable, that is, non-stoquastic<sup>13</sup>. Such Hamiltonians would enable universal adiabatic quantum computing<sup>14,15</sup> and improve the performance for difficult instances of classical optimization problems<sup>16</sup>. Additionally, simulating interacting fermions for applications in physics and chemistry requires non-stoquastic Hamiltonians<sup>1,17</sup>. In general, these Hamiltonians are more difficult to study classically, because Monte Carlo simulations fail to converge owing to the ‘sign problem’<sup>18</sup>. A hallmark of non-stoquastic Hamiltonians is the need for several distinct types of coupling, for example,  $\sigma_z\sigma_z$  and  $\sigma_x\sigma_x$  couplings with different signs, where  $\sigma_x$  and  $\sigma_z$  are Pauli operators. With a digitized approach, different couplings can be constructed without change of hardware. Long-range multibody interactions can be assembled to aid in quantum tunnelling<sup>19</sup> or to encode the non-local terms for fermionic simulations<sup>20,21</sup>. And finally, analogue systems exhibit noise, which can thwart the evolution, whereas digital systems can be fully fault-tolerant. Crucially, this ability makes the digitized approach scalable, because any non-corrected implementation is ultimately limited by the accumulation of error. Our experiment addresses the

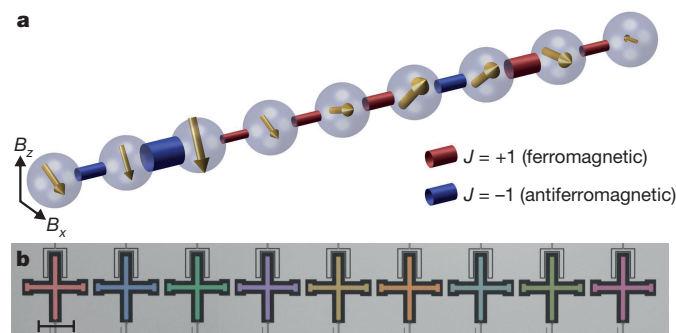
challenge of adiabatically evolving to final problem Hamiltonians that are non-stoquastic.

We explore the adiabatic quantum evolutions of one-dimensional spin chains with nearest-neighbour coupling. We start with a simple ferromagnetic problem to visualize the adiabatic evolution process. We identify specific error contributions, and follow up by exploring the scaling of errors with system size. We finish by testing the device on random stoquastic and non-stoquastic problems. The initial (‘I’) and problem (‘P’) Hamiltonians are

$$H_I = -B_{x,I} \sum_i \sigma_x^i$$

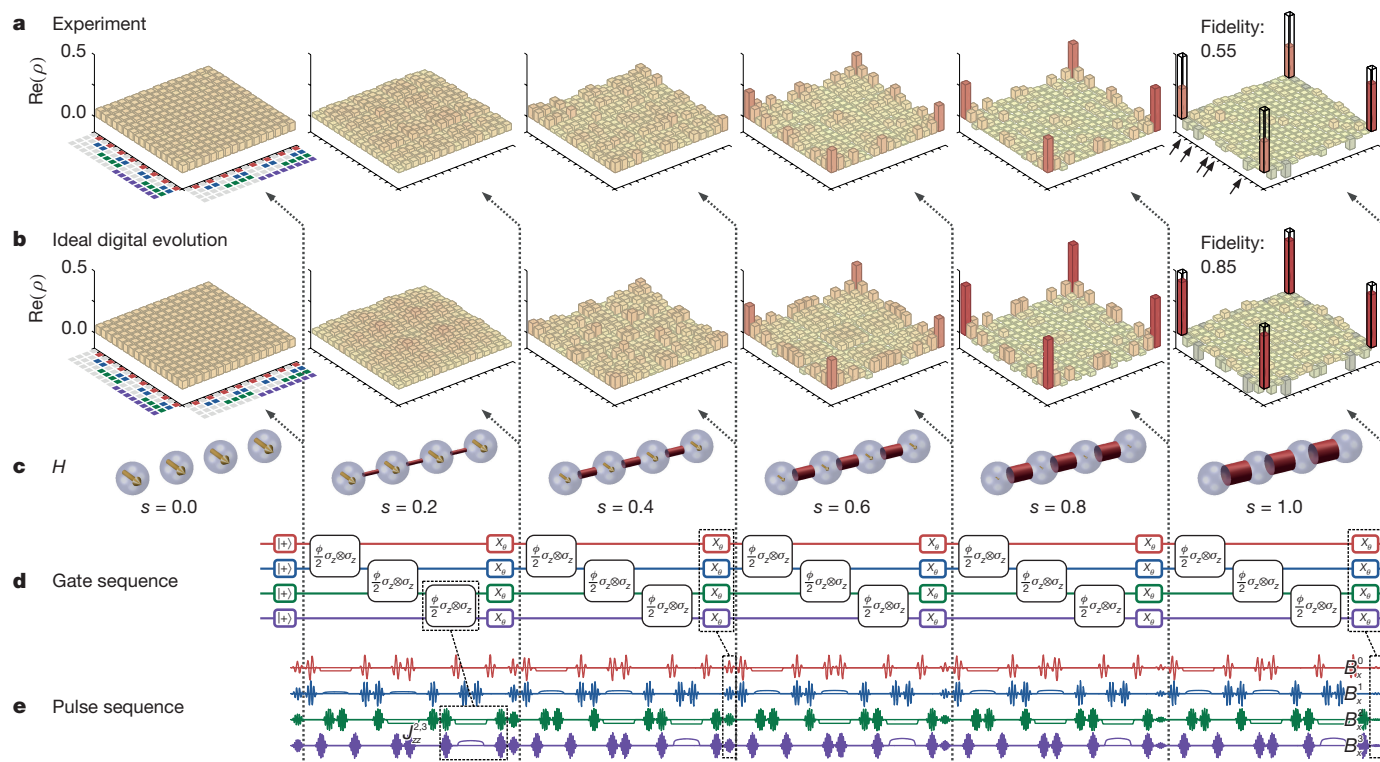
$$H_P = - \sum_i (B_z^i \sigma_z^i + B_x^i \sigma_x^i) - \sum_i (J_{zz}^{i,i+1} \sigma_z^i \sigma_z^{i+1} + J_{xx}^{i,i+1} \sigma_x^i \sigma_x^{i+1})$$

where  $B_z^i$  and  $B_x^i$  denote local field strengths of the  $i$ th qubit,  $J_{zz}^{i,i+1}$  and  $J_{xx}^{i,i+1}$  denote the  $\sigma_z\sigma_z$  and  $\sigma_x\sigma_x$  coupling strengths, respectively, between qubits  $i$  and  $i+1$  and  $B_{x,I}$  denotes the initial field strength, which is equal for all qubits. The Ising model is recovered when  $B_x = J_{xx} = 0$  for all  $i$ . We initialize the system with  $H_I$  and vary the system Hamiltonian to the final problem:  $H = sH_P + (1-s)H_I$ , with  $s$  going from 0 to 1. An example problem is shown in Fig. 1a.



**Figure 1 | Spin-chain problem and device.** **a**, We implement one-dimensional spin problems with variable local fields and couplings between adjacent spins. An example of a stoquastic problem Hamiltonian with local  $x$  and  $z$  fields, indicated by the gold arrows in the spheres, and  $\sigma_z\sigma_z$  couplings, whose strength is indicated by the radius of the links, is shown. Red denotes a ferromagnetic ( $J = +1$ ) and blue an antiferromagnetic ( $J = -1$ ) link. The problem Hamiltonian is for the instance shown in Fig. 4c. **b**, Optical picture of the superconducting quantum device with nine Xmon<sup>22</sup> qubits  $Q_0$ – $Q_8$  (false-coloured cross-shaped structures), made from aluminium (light) on a sapphire substrate (dark). Connections to read-out resonators are at the top; control wiring is at the bottom. Scale bar, 200  $\mu\text{m}$ .

<sup>1</sup>Google Inc., Santa Barbara, California 93117, USA. <sup>2</sup>Google Inc., Venice, California 90291, USA. <sup>3</sup>Department of Physical Chemistry, University of the Basque Country UPV/EHU, Apartado 644, E-48080 Bilbao, Spain. <sup>4</sup>Department of Physics, University of California, Santa Barbara, California 93106, USA. <sup>5</sup>KERBASQUE, Basque Foundation for Science, Maria Diaz de Haro 3, 48013 Bilbao, Spain. <sup>†</sup>Present address: IBM T. J. Watson Research Center, Yorktown Heights, New York 10598, USA.



**Figure 2 | Quantum state tomography of the digital evolution into a Greenberger–Horne–Zeilinger state.** A four-qubit system is adiabatically evolved from an initial Hamiltonian in which all spins are aligned along the  $x$  axis to a problem Hamiltonian with equal ferromagnetic couplings between adjacent qubits ( $J_{zz} = 2$ ). **a**, Real part of the experimental density matrix  $\rho$  at the start (left-most panel) and after each Trotter step, showing the growth of the major elements on the four corners, measured using quantum state tomography. The target state is shown with a black outline in the right-most panel. The final state has a fidelity of 0.55. Coloured squares surrounding the left-most panel indicate qubit indices: for example,  $Q_0$  being excited is indicated by a red square. Black arrows indicate notable elements for states that differ from the target state by a single kink. **b**, As in **a**, but for the ideal digitized evolution,

showing major elements on the four corners as well as other populations and correlations. **c**, Hamiltonian at different  $s$ , showing the vanishing transversal field and increasing coupling strength; arrows and links as in Fig. 1a. **d**, Gate sequence showing initialization and the five Trotter steps. **e**, Pulse sequence, showing the single-qubit microwave gates (wave-like pulses) and frequency detuning (rectangular-like) pulses. Corresponding interactions and local field terms are highlighted. The displayed five-step algorithm is 2.1- $\mu$ s long. Colours correspond to the physical qubits in Fig. 1b. Implementations of  $\sigma_z\sigma_z$  coupling and local  $x$ -fields are highlighted. Angles of rotation are denoted by  $\phi$  and  $\theta$ . See Supplementary Information for imaginary parts of the density matrices and the ideal continuous evolution.

The spin system is formed by a superconducting circuit with nine qubits. The qubits are the cross-shaped structures<sup>22</sup>, patterned out of an aluminium layer on top of a sapphire substrate, and arranged in a linear chain; see Fig. 1b. Each qubit is capacitively coupled to its nearest neighbours, and can be individually controlled and measured; for details see ref. 23. By tuning the frequencies of the qubits we can implement a tunable controlled-phase entangling gate. We use the first-order Trotter expansion to digitize<sup>24</sup>. The evolution is divided into many steps and implemented using gates; see Supplementary Information.

For quantifying digitized adiabatic evolutions there are four sets of data: (1) the ideal continuous time evolution, for infinite time, which is free of error and provides the perfect solution, and which we refer to as the ‘target state’; (2) the ideal continuous time evolution for a finite time  $T$ , which is sensitive to non-adiabatic errors, and which we call ‘ideal continuous evolution’; (3) the ‘ideal digital evolution’, where the finite ideal continuous evolution is digitized, and which therefore includes digital error as well as non-adiabatic errors; and (4) the experimental results, which include a contribution from gate errors as well.

We start with a ferromagnetic chain problem with  $N = 4$  spins, and equal coupling strength  $J_{zz} = 2$ . The qubits are initialized in the  $|+\rangle^{\otimes N}$  state, and we use five steps to evolve the system to the problem Hamiltonian, performing quantum state tomography after each step. We linearly decrease the  $B_x$  term to zero, starting at  $B_x = 2$ , and simultaneously increase the coupling strength from 0 to 2, ending the evolution at a scaled time of  $|J|T = 6$ . The density matrices are shown in Fig. 2a. With each step, the quantum state evolves and matrix elements

in the middle vanish while the elements at the four corners grow to form the density matrix  $\rho$  of the Greenberger–Horne–Zeilinger (GHZ) state—the solution to the ferromagnetic problem—with a fidelity  $\text{tr}(\rho_{\text{target-state}}\rho) = 0.55$ . The density matrix is constrained to be physical<sup>25</sup>. The ideal digital evolution is plotted in Fig. 2b, reaching a fidelity of 0.85. The Hamiltonian during evolution, construction of the algorithm and the pulse sequence are shown in Fig. 2c–e. In each Trotter step, we perform a  $\sigma_z\sigma_z$  operation on each pair to implement the ferromagnetic  $\sigma_z\sigma_z$  coupling, followed by single-qubit rotations around the  $x$  axis to simulate the transversal magnetic field. In the pulse sequence, the rectangular-like frequency detuning pulses indicate where  $\sigma_z\sigma_z$  interaction is implemented by bringing qubits near resonance (highlighted for  $s = 0.2$  in Fig. 2d, e). The wave-like pulses are microwave gates. The decrease in  $B_x$  is reflected by the reduction in amplitudes of the corresponding pulses (highlighted for  $s = 0.4$  and  $s = 1.0$  in Fig. 2d, e). Additional microwave echo pulses decrease coupling to other qubits and the environment. We find mean phase errors from neighbouring parasitic interactions to be around 0.05 rad, equivalent to an error contribution below  $10^{-3}$  (see Supplementary Information).

The experiment in Fig. 2 shows that digital synthesis of adiabatic evolutions can successfully be implemented in a solid-state quantum platform. Using five Trotter steps, 15 entangling gates and 144 single-qubit microwave gates, the system produces a GHZ state with a fidelity that indicates genuine entanglement. It shows that complex pulse sequences are possible, and that the errors make sense: the fidelity of the experimental data with respect to the ideal digital evolution is



0.64. The overlap between the ideal digital evolution and ideal continuous time evolution for finite time is 0.93, and the overlap of this continuous evolution with the GHZ state (see Supplementary Information) is 0.88. The product of the above three values (0.52) is close to the experimental fidelity of 0.55, and shows that the experimental error is a combination of non-adiabatic, digitization and gate errors. Adopting the entangling gate error of  $7.4 \times 10^{-3}$  and  $8 \times 10^{-4}$  as measured in ref. 25, we expect an accumulated gate error of 0.23 whereas we find an infidelity of 0.36; we attribute the difference to errors in maintaining the phases of the four-qubit system for a duration of  $2.1 \mu\text{s}$ .

An important feature of the errors is the prevalence of populations and correlations of the  $|0001\rangle$ ,  $|0011\rangle$  and  $|0111\rangle$  states and their bit-wise inverses; see arrows in Fig. 2a. Their elements are also present in the ideal digital results and in the ideal continuous evolutions (see Supplementary Information). These are states that deviate by a single kink from the target state, having a residual energy of  $2|J|$ , indicating the presence of non-adiabatic errors. These kink errors are connected to the formation of defects during a phase transition, as described by the Kibble–Zurek mechanism<sup>26,27</sup>.

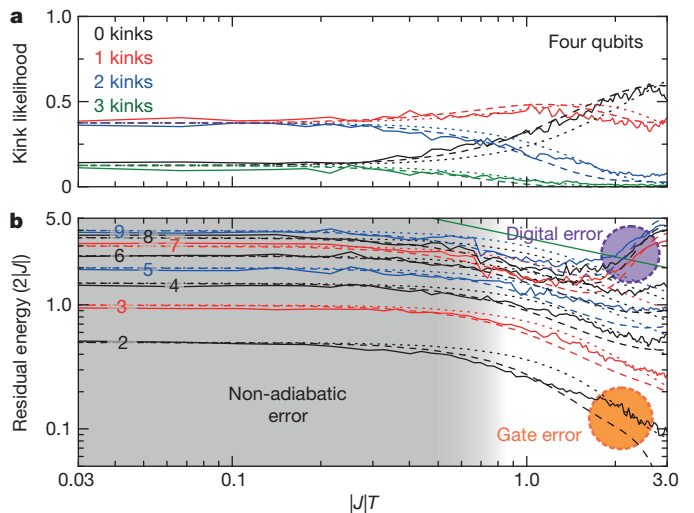
To explore the scaling of errors we vary the system size from two to nine qubits and measure the likelihood of kinks and residual energy. We keep the ferromagnetic problem Hamiltonian,  $J_{zz} = 2$ , but vary the scaled time such that  $|J|T$  goes from 0 to 3. For the two- to six-qubit systems we use five Trotter steps and for seven to nine qubits we use two steps, to limit the total number of gates. The kink likelihood for the four-qubit system is shown in Fig. 3a. Here, the likelihood of one kink is given by the sum of the probabilities of all states with one kink. When increasing  $|J|T$  from 0 to 3 the kink likelihood decreases, and the likelihood of no kinks increases (black line in Fig. 3a). The experimental data closely follow the ideal digital evolution (dashed lines in Fig. 3a). This picture is repeated for all systems; see Supplementary Information.

The kink likelihood indicates that the final state has residual energy, because a state with a single kink has energy  $2|J|$  above the target state. The residual energies for all systems are plotted in Fig. 3b. Initially, the residual energy is constant at  $|J|T \approx 0$ , then starts to decrease around  $|J|T \approx 0.5$ , following the ideal digital (dashed lines in Fig. 3b) and ideal continuous time (dotted lines in Fig. 3b) evolutions. For two to six qubits, this decrease continues until the traces start to settle around  $|J|T = 3$ . For the seven- to nine-qubit systems, the residual energy starts to increase again around  $|J|T = 2$ , following the ideal digital evolution. See Supplementary Information for the pulse sequence for the nine-qubit experiment, all kink likelihoods and the differences between the residual energies.

The main result is that Fig. 3 distinctly shows the different contributions to error (highlighted): for  $|J|T \ll 1$ , the residual energy is dominated by non-adiabatic errors because the evolution moves too fast. For  $|J|T > 2$ , the flattening out of the residual energy for the configurations with two to six qubits indicates that gate errors dominate, because the predictions from the ideal digital evolutions are substantially lower. For the larger qubit configurations with seven to nine qubits, the residual energy follows the digital predictions upwards, indicating that digitization errors dominate. In addition, the residual energy visibly decreases at  $|J|T = 1$  for all configurations, implying that the digitized evolutions are able to approach the target state even for nine qubits.

We also applied local fields to explore the lifting of degeneracy and generation of long-range correlations; see Supplementary Information.

We next discuss how the digitized approach can solve stoquastic and non-stoquastic problems with comparable performance, by testing random problems on three, six, seven, eight and nine qubits. Problems have local fields and couplings with random strength and sign. We independently choose  $B_z$  and  $B_x$  from  $[-2, 2]$  for each spin and  $J_{zz}$  from  $[-2, -0.5]$  or  $[0.5, 2]$  for each link. This creates a random Ising problem with frustration. For non-stoquastic problems we also add  $J_{xx}$  coupling for each link, with values from  $[-2, -0.5]$  or  $[0.5, 2]$ , effectively doubling the amount of entangling gates. We avoid small couplings to reduce the number of gates. For the three-qubit systems we used



**Figure 3 | Kink errors, residual energy and scaling with system size.**

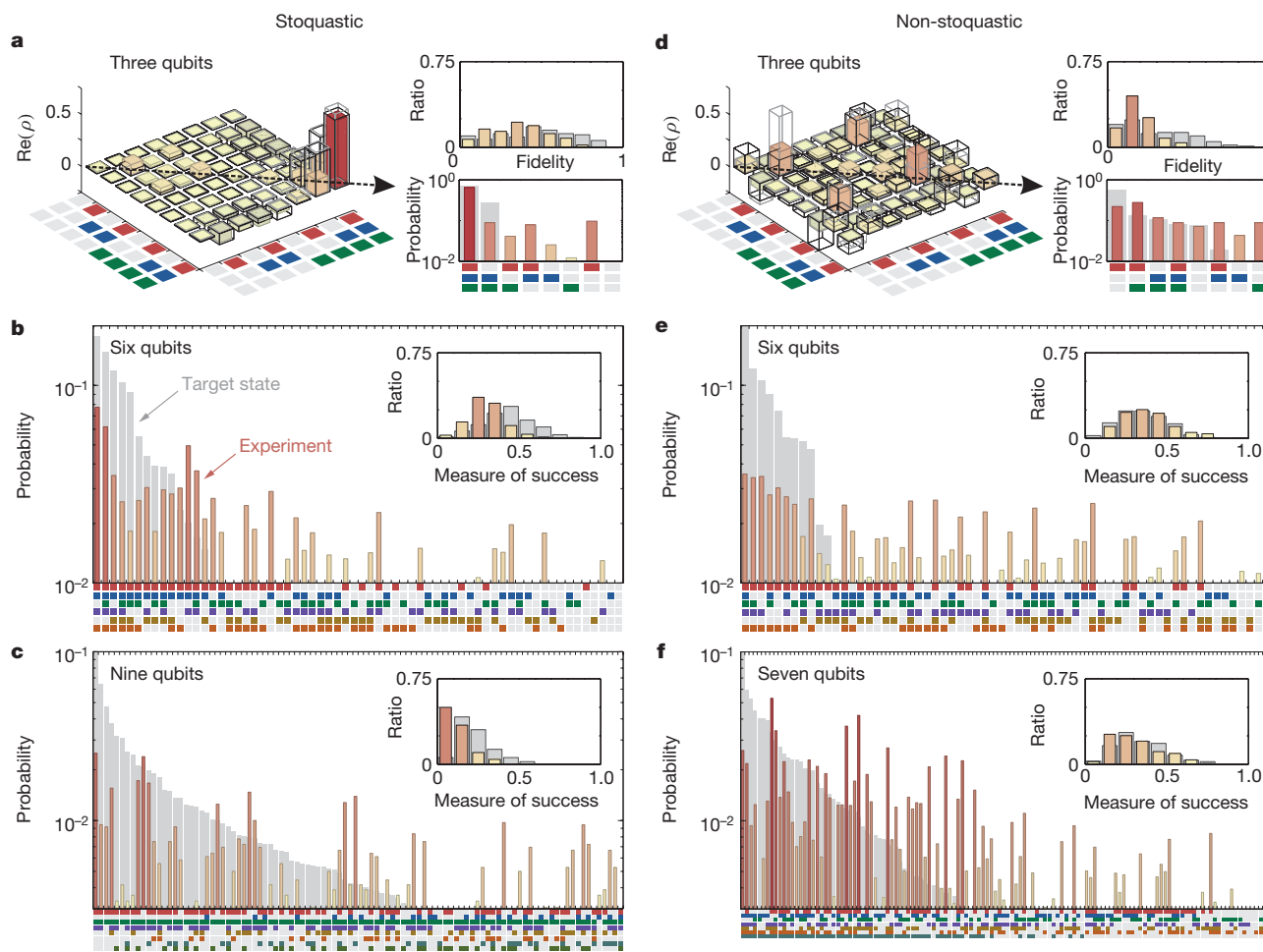
**a**, Kink likelihood for the four-qubit configuration. Solid lines, experiment; dashed lines, ideal digital evolution; dotted lines, ideal continuous time evolution. **b**, Residual energy in the adiabatic evolutions of ferromagnetic chains ( $J_{zz} = 2$ ) in configurations with two to nine qubits (as indicated by the coloured-coded numerals). The green solid line shows the ideal square-root trend for the large-scale limit (Supplementary Information). Distinct contributions to error are highlighted.

quantum state tomography on 100 separate instances to include off-diagonal elements in the fidelity metrics. For six or more qubits tomography is not practical and so we measured the correlated probabilities on 250 separate instances, and use a measure of success that is equal to  $|\langle \Psi_{\text{ideal}} | \Psi \rangle|^2$  (with  $\Psi$  the wavefunction) to first order and sets an upper bound on the fidelity:  $(\sum_k \sqrt{P_{k,\text{ideal}} P_k})^2$ , in which  $P_{k,\text{ideal}}$  and  $P_k$  are probabilities and  $k$  runs over the computational basis. In Fig. 4 we show the results for stoquastic problems with three, six and nine spins, and non-stoquastic problems with three, six and seven spins. For each case, we highlight a single instance and show histograms of the fidelities.

For the three-spin stoquastic problems, the real part of the density matrix of one instance and a histogram of its diagonal elements are shown in Fig. 4a. In the tomography plot (left panel of Fig. 4a), we overlay the experimental results (colour) with the ideal digital (black) and target state (grey) results. For this example, we find fidelities  $\text{tr}(\rho_{\text{ideal-digital}}\rho) = 0.70$  and  $\text{tr}(\rho_{\text{target-state}}\rho) = 0.63$ . In the top right panel of Fig. 4a, we show the histograms for all instances of the fidelities  $\text{tr}(\rho_{\text{target-state}}\rho)$  in colour. The fidelity of the ideal digital evolution with respect to the target state is shown in grey. Stoquastic problems with six and nine qubits are displayed in Fig. 4b and c, respectively. The main figures show the measured probabilities (colour) sorted by the target state results (grey), and the insets display the histograms. Results for the non-stoquastic problems are displayed in Fig. 4d–f.

The key result from Fig. 4 is that the physical system can find solutions to non-stoquastic problems with a performance similar to that of stoquastic problems. The three-qubit examples show major diagonal as well as off-diagonal elements close to the expected positions. For six and more qubits, the coloured bars in the example instances are mostly on the left, indicating that the system has a clear preference for returning the probabilities associated with the ideal solutions.

The physical system produces results that are comparable to the expectations, as demonstrated by the histograms showing a substantial overlap between experiment and theory. Moreover, the numbers are consistent, as we now discuss for the six-qubit stoquastic example. The mean success rate between the ideal adiabatic evolution and target state is  $0.59 \pm 0.01$ , indicating that the scaled time is large enough to capture the evolution dynamics. The mean success rate of the ideal digitized evolution with respect to the ideal adiabatic evolution is  $0.73 \pm 0.01$ , indicating a proper Trotterization of the evolution. Finally, the value for



**Figure 4 | Digital evolutions of random stoquastic and non-stoquastic problems.** As stoquastic problems we use frustrated Ising Hamiltonians, with random local  $x$  and  $z$  fields, and random  $\sigma_x\sigma_z$  couplings. **a–c**, Stoquastic results for three, six and nine qubits. **a**, For three qubits we have done tomography. An example instance is provided on the left, where we show the real part of the density matrix  $\rho$ . Coloured bars denote the experimental data, and black and grey outlined bars show the ideal digital evolution and the target state, respectively. The diagonals of the experiment (colour) and the target state (grey) are shown in the bottom right panel (as indicated by the dashed arrow), sorted by ideal target state results. The fidelity results for all 100 instances are summarized in the histogram (top right), where ratio denotes the normalized occurrence; coloured bars, fidelities of experimental results with respect to the target

state; grey bars, fidelities of the ideal digital evolution with respect to the target state. **b, c**, The correlated probabilities for six (**b**) and nine (**c**) qubits, sorted by target state results. Experimental data are in colour, the target state is in grey. The results for all 250 instances are summarized in the insets. For the nine-qubit instance (**c**), the first 100 elements are shown. In **a–c**, the coloured squares surrounding or below the plots indicate qubit indices, as in Fig. 2. **d–f**, As in **a–c**, but for non-stoquastic problems, which have additional random  $\sigma_x\sigma_x$  couplings. Here we plot the data for three, six and seven qubits, for which the average measure of success is above the random baseline (not shown; see text). The results show that the system can find the ground states of stoquastic and non-stoquastic Hamiltonians with similar performance.

the experimental evolution with respect to the ideal digitized evolution is  $0.714 \pm 0.006$ , indicating that the experiment follows the ideal digital evolution reasonably well. The product of these three numbers, 0.31, is very close to the mean value between the experimental data and the target state,  $0.296 \pm 0.007$ . This shows that the experimental errors arise from comparable contributions of non-adiabatic, digital and gate errors. For the six-qubit non-stoquastic case, experimental-to-target state values are higher than this product, suggesting that errors partially cancel. A further reason for the higher success rate could be that the presence of  $\sigma_x\sigma_x$  terms is helpful for difficult problems in general<sup>16</sup>. This experiment took up to nine qubits and up to  $10^3$  gates. See Supplementary Information for pulse sequences, gate counts, problem parameters and additional metrics.

To further quantify the performance of the system, we compare experimental and random probabilities with the theoretical results. In essence, we take a uniform random distribution over the  $2^N$  possible measurement outputs as a baseline sanity check. We find that, for the stoquastic problems, the measures of success of all six- to nine-qubit

configurations are significantly above this baseline: for six qubits, the success measure of the experimental data with respect to the target state is  $0.296 \pm 0.007$ , whereas using uniform random probabilities produces a value of  $0.168 \pm 0.005$ . For the nine-qubit case the numbers are  $0.122 \pm 0.006$  for the experimental data and  $0.074 \pm 0.004$  for random. For the non-stoquastic problems the numbers are  $0.380 \pm 0.009$  and  $0.335 \pm 0.008$  for the six-qubit configuration, and  $0.311 \pm 0.009$  and  $0.277 \pm 0.008$  for the seven-qubit configuration. A complete listing for all configurations is provided in Supplementary Information.

This experiment shows that digital synthesis of the adiabatic evolutions can be used to find signatures of the ground states of random stoquastic and non-stoquastic problems. Errors arise from a comparable contribution of non-adiabatic, digital and gate errors, and success rates are significantly above a uniform random baseline. For larger qubit systems, the number of Trotter steps needs to be limited to reduce the accumulation of gate error, in turn limiting the evolution we can simulate. Therefore, the experimental error is larger, arising from a

combination of gate, digitization and non-adiabatic error. However, in an error-corrected system, the number of gates is in principle unconstrained, digitization can be made arbitrarily accurate and one can move more slowly through critical parts of the evolution. Although we have used Trotterization<sup>28</sup>, the scaling of the digitization becomes more appealing with recent methods based on the truncation of Taylor series<sup>29</sup>. See Supplementary Information for further motivations and discussions.

We believe that the digitized approach to adiabatic quantum evolutions of complex problems—where local fields, variable coupling strengths and types, and multibody interactions can be constructed—would become viable on the small scale with lower gate errors, and that large-scale applications could be achieved in conjunction with error correction. We hope our work accelerates further improvements in superconducting quantum systems and motivates research into the encoding and measurement of non-stoquastic computational problems. In addition, we anticipate that these results encourage work on the efficient digitization of algorithms for small- and large-scale systems, for which reducing the effects of noise by, for example, dynamical decoupling techniques, or reducing the circuit complexity is paramount.

Received 13 November 2015; accepted 1 March 2016.

- Feynman, R. P. Simulating physics with computers. *Int. J. Theor. Phys.* **21**, 467–488 (1982).
- Aspuru-Guzik, A., Dutoi, A. D., Love, P. J. & Head-Gordon, M. Simulated quantum computation of molecular energies. *Science* **309**, 1704–1707 (2005).
- Farhi, E., Goldstone, J., Gutmann, S. & Sipser, M. Quantum computation by adiabatic evolution. Preprint at <http://arxiv.org/abs/quant-ph/0001106> (2000).
- Farhi, E. *et al.* A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem. *Science* **292**, 472–475 (2001).
- Nishimori, H. *Statistical Physics of Spin Glasses and Information Processing: An Introduction* (Oxford Univ. Press, 2001).
- Lloyd, S. Universal quantum simulators. *Science* **273**, 1073–1078 (1996).
- Bravyi, S. B. & Kitaev, A. Yu. Quantum codes on a lattice with boundary. Preprint at <http://arxiv.org/abs/quant-ph/9811052> (1998).
- Fowler, A. G., Mariantoni, M., Martinis, J. M. & Cleland, A. N. Surface codes: towards practical large-scale quantum computation. *Phys. Rev. A* **86**, 032324 (2012).
- Steffen, M. *et al.* Experimental implementation of an adiabatic quantum optimization algorithm. *Phys. Rev. Lett.* **90**, 067903 (2003).
- Lanyon, B. P. *et al.* Universal digital quantum simulation with trapped ions. *Science* **334**, 57–61 (2011).
- Barends, R. *et al.* Digital quantum simulation of fermionic models with a superconducting circuit. *Nat. Commun.* **6**, 7654 (2015).
- Salathé, Y. *et al.* Digital quantum simulation of spin models with circuit quantum electrodynamics. *Phys. Rev. X* **5**, 021027 (2015).
- Bravyi, S., DiVincenzo, D. P., Oliveira, R. I. & Terhal, B. M. The complexity of stoquastic local Hamiltonian problems. *Quantum Inf. Comput.* **8**, 361–385 (2008).
- Aharonov, D. *et al.* Adiabatic quantum computation is equivalent to standard quantum computation. *SIAM Rev.* **50**, 755–787 (2008).
- Lloyd, S. & Terhal, B. M. Adiabatic and Hamiltonian computing on a 2D lattice with simple two-qubit interactions. *New J. Phys.* **18**, 023042 (2016).
- Crosson, E., Farhi, E., Lin, C. Y.-Y., Lin, H.-H. & Shor, P. Different strategies for optimization using the quantum adiabatic algorithm. Preprint at <http://arxiv.org/abs/1401.7320> (2014).
- Babbush, R., Love, P. J. & Aspuru-Guzik, A. Adiabatic quantum simulation of quantum chemistry. *Sci. Rep.* **4**, 6603 (2014).
- Troyer, M. & Wiese, U.-J. Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations. *Phys. Rev. Lett.* **94**, 170201 (2005).
- Boixo, S. *et al.* Computational multiqubit tunnelling in programmable quantum annealers. *Nat. Commun.* **7**, 10327 (2016).
- Bravyi, S. B. & Kitaev, A. Yu. Fermionic quantum computation. *Ann. Phys.* **298**, 210–226 (2002).
- Seeley, J. T., Richard, M. J. & Love, P. J. The Bravyi–Kitaev transformation for quantum computation of electronic structure. *J. Chem. Phys.* **137**, 224109 (2012).
- Barends, R. *et al.* Coherent Josephson qubit suitable for scalable quantum integrated circuits. *Phys. Rev. Lett.* **111**, 080502 (2013).
- Kelly, J. *et al.* State preservation by repetitive error detection in a superconducting quantum circuit. *Nature* **519**, 66–69 (2015).
- Suzuki, M. Fractal decomposition of exponential operators with applications to many-body theories and Monte Carlo simulations. *Phys. Lett. A* **146**, 319–323 (1990).
- Barends, R. *et al.* Superconducting quantum circuits at the surface code threshold for fault tolerance. *Nature* **508**, 500–503 (2014).
- Kibble, T. W. B. Some implications of a cosmological phase transition. *Phys. Rep.* **67**, 183–199 (1980).
- Zurek, W. H. Cosmological experiments in superfluid helium? *Nature* **317**, 505–508 (1985).
- Wiebe, N., Berry, D., Hoyer, P. & Sanders, B. C. Higher order decompositions of ordered operator exponentials. *J. Phys. A* **43**, 065203 (2010).
- Berry, D., Childs, A. M., Cleve, R., Kothari, R. & Somma, R. D. Simulating Hamiltonian dynamics with a truncated Taylor series. *Phys. Rev. Lett.* **114**, 090502 (2015).

Supplementary Information is available in the online version of the paper.

**Acknowledgements** We acknowledge support from Spanish MINECO FIS2012-36673-C03-02; Ramón y Cajal grant RYC-2012-11391; UPV/EHU UFI 11/55 and EHUA14/04; Basque Government IT472-10; a UPV/EHU PhD grant; and PROMISCE and SCALEQIT EU projects. Devices were made at the UC Santa Barbara Nanofabrication Facility, a part of the NSF-funded National Nanotechnology Infrastructure Network, and at the NanoStructures Cleanroom Facility.

**Author Contributions** R. Barends, A.S. and L.L. designed the experiment, with E.S., H.N. and J.M.M. providing supervision and A. Mezzacapo, U.L.H. and R. Babbush providing additional theoretical support. R. Barends, A.S., L.L. and R. Babbush co-wrote the manuscript with E.S., H.N. and J.M.M. R. Barends, A.S. and L.L. performed the experiment and analysed the data. The device was designed by R. Barends and J.K. All authors contributed to the fabrication process, experimental set-up and manuscript preparation.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R. Barends ([barends@google.com](mailto:barends@google.com)) or A.S. ([shabani@google.com](mailto:shabani@google.com)).



# Metastable high-entropy dual-phase alloys overcome the strength-ductility trade-off

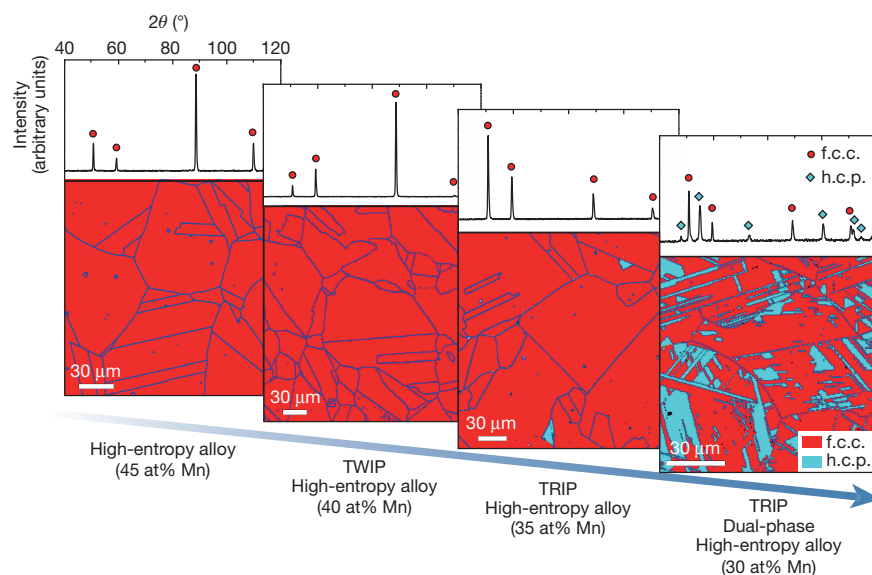
Zhiming Li<sup>1</sup>, Konda Gokuldoss Pradeep<sup>1</sup>, Yun Deng<sup>1</sup>, Dierk Raabe<sup>1</sup> & Cemal Cem Tasan<sup>1,2</sup>

Metals have been mankind's most essential materials for thousands of years; however, their use is affected by ecological and economical concerns. Alloys with higher strength and ductility could alleviate some of these concerns by reducing weight and improving energy efficiency. However, most metallurgical mechanisms for increasing strength lead to ductility loss, an effect referred to as the strength-ductility trade-off<sup>1,2</sup>. Here we present a metastability-engineering strategy in which we design nanostructured, bulk high-entropy alloys with multiple compositionally equivalent high-entropy phases. High-entropy alloys were originally proposed to benefit from phase stabilization through entropy maximization<sup>3–6</sup>. Yet here, motivated by recent work that relaxes the strict restrictions on high-entropy alloy compositions by demonstrating the weakness of this connection<sup>7–11</sup>, the concept is overturned. We decrease phase stability to achieve two key benefits: interface hardening due to a dual-phase microstructure (resulting from reduced thermal stability of the high-temperature phase<sup>12</sup>); and transformation-induced hardening (resulting from the reduced mechanical stability of the room-temperature phase<sup>13</sup>). This combines the best of two worlds: extensive hardening due to the decreased phase stability known from advanced steels<sup>14,15</sup> and massive solid-solution strengthening of high-entropy alloys<sup>3</sup>. In our transformation-induced plasticity-assisted, dual-phase high-entropy alloy (TRIP-DP-HEA), these two contributions lead respectively to enhanced trans-grain and inter-grain slip resistance, and hence, increased strength. Moreover, the increased strain hardening capacity that is enabled by dislocation hardening of the stable phase and transformation-induced hardening of the metastable phase produces increased ductility. This combined increase in strength

and ductility distinguishes the TRIP-DP-HEA alloy from other recently developed structural materials<sup>16,17</sup>. This metastability-engineering strategy should thus usefully guide design in the near-infinite compositional space of high-entropy alloys.

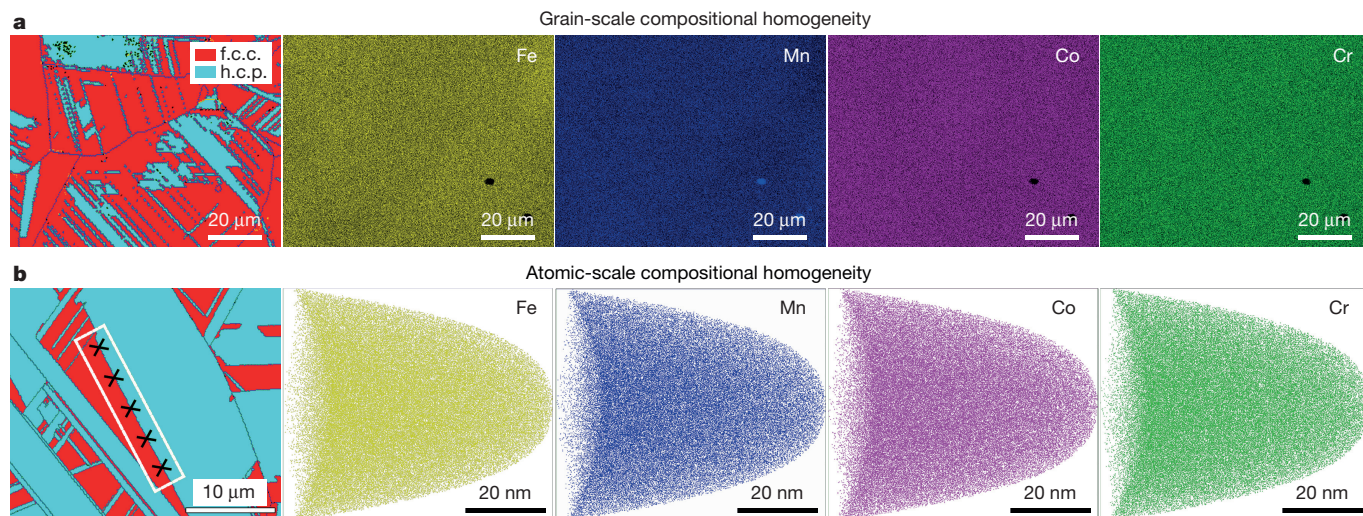
To realize the TRIP-DP-HEA concept, we switch from the equiatomic  $\text{Fe}_{20}\text{Mn}_{20}\text{Ni}_{20}\text{Co}_{20}\text{Cr}_{20}$  (atomic per cent, at%)<sup>6</sup> system to the non-equiatomic  $\text{Fe}_{80-x}\text{Mn}_x\text{Co}_{10}\text{Cr}_{10}$  (at%) system, which exhibits partial martensitic transformation of the face-centred cubic (f.c.c.) to the hexagonal close-packed (h.c.p.) phase upon cooling from the high-temperature single-phase region. This change enables development of a dual-phase microstructure in which both phases obtain the maximum benefit of the solid-solution strengthening effect and one phase, owing to the decreased stacking fault energy<sup>18</sup>, undergoes deformation-induced displacive transformation. The partial martensitic transformation during quenching is the only possible approach that can lead to the formation of a DP-HEA with phases of identical chemical composition (that is, high-entropy phases). The alloys were synthesized with varying Mn contents in a vacuum induction furnace using pure metals, hot-rolled to 50% thickness at 900 °C, homogenized at 1,200 °C for 2 h in an Ar atmosphere, and water-quenched. Further grain refinement was achieved by cold-rolling (to 60% thickness) and 3-min annealing at 900 °C in an Ar atmosphere. The chemical composition of the HEAs measured by wet-chemical analysis is given in Extended Data Table 1.

Microstructure characterization down to 30-nm resolution reveals that the  $\text{Fe}_{80-x}\text{Mn}_x\text{Co}_{10}\text{Cr}_{10}$  (at%) system indeed demonstrates the targeted change in phase stability (see the X-ray diffraction (XRD) and electron backscatter diffraction (EBSD) data in Fig. 1). A single f.c.c. phase structure was obtained when the Mn content was 45 at% and 40 at% ( $\text{Fe}_{35}\text{Mn}_{45}\text{Co}_{10}\text{Cr}_{10}$  and  $\text{Fe}_{40}\text{Mn}_{40}\text{Co}_{10}\text{Cr}_{10}$ , respectively).



**Figure 1 | XRD patterns and EBSD phase maps of  $\text{Fe}_{80-x}\text{Mn}_x\text{Co}_{10}\text{Cr}_{10}$  ( $x = 45$  at%, 40 at%, 35 at% and 30 at%) HEAs.**  $\theta$  is the Bragg angle. The Mn content plays an important part in phase constitution, tuning phase stability for the activation of specific displacing transformation mechanisms, for example, enabling TWIP or TRIP effects. We note that the 35 at% Mn alloy has only trace amounts of the h.c.p. phase, and hence is not referred to as a DP-HEA.

<sup>1</sup>Max-Planck-Institut für Eisenforschung, Max-Planck-Straße 1, 40237 Düsseldorf, Germany. <sup>2</sup>Department of Materials Science and Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139 USA.



**Figure 2 | Elemental homogeneity among the two phases of  $\text{Fe}_{50}\text{Mn}_{30}\text{Co}_{10}\text{Cr}_{10}$  (at%) HEA.** **a**, Energy-dispersive spectroscopy maps of the shown EBSD-mapped sample region. **b**, Three-dimensional APT tip reconstructions of Fe, Mn, Co, Cr atom positions in a typical APT tip from the EBSD-mapped phase boundary. The crosses refer to the positions that the APT tips were taken from.

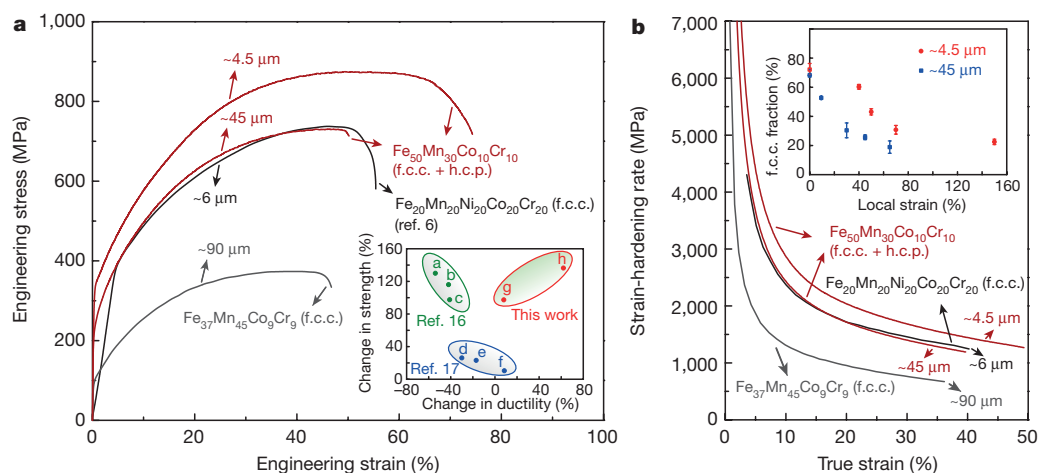
These two alloys demonstrate a transition in the deformation mechanisms from dislocation-dominated plasticity in the former<sup>8</sup> to twinning-induced plasticity (TWIP) in the latter<sup>10</sup>, confirming the targeted stability trend realized by tuning the stacking fault energy. A further decrease to 35 at% Mn leads to traces of h.c.p. phase (not captured by XRD). Finally, a decrease to 30 at% Mn ( $\text{Fe}_{50}\text{Mn}_{30}\text{Co}_{10}\text{Cr}_{10}$ ) successfully produces the desired dual-phase microstructure with  $\sim 28\%$  h.c.p. phase. This alloy is analysed in more detail in the following.

The two phases constituting the as-quenched  $\text{Fe}_{50}\text{Mn}_{30}\text{Co}_{10}\text{Cr}_{10}$  (at%) alloy are the f.c.c.  $\gamma$  matrix (of  $\sim 45\text{-}\mu\text{m}$  grain size) and the h.c.p.  $\epsilon$  phase laminate layers (ranging from several nanometres to  $10\text{-}\mu\text{m}$  in thickness). In Fig. 2, energy dispersive spectroscopy and atom probe tomography (APT) maps are also provided for the corresponding EBSD maps, respectively, to reveal the compositional distribution among the two phases. The energy dispersive spectroscopy maps in Fig. 2a show that all elements are uniformly distributed, suggesting that both phases benefit from the same level of solid solution strengthening. APT tips were lifted out from a phase boundary region (using the method outlined in ref. 19) shown in the EBSD phase map in Fig. 2b, to

rule out the possibility of atomic-scale elemental partitioning between the f.c.c. and the h.c.p. phases. The analysis reveals that the investigated volume has an overall composition of  $\text{Fe}_{48.6}\text{Mn}_{27.6}\text{Co}_{11.3}\text{Cr}_{12.3}$  (at%), showing values near the nominal bulk composition. No apparent elemental segregations can be observed in the three-dimensional reconstructions (Fig. 2b) or from the statistical binomial frequency distribution analyses (Extended Data Fig. 1), confirming the uniform distribution of all elements even at phase boundaries. This is different from Mn-containing steels, which show substantial chemical gradients across phase boundaries<sup>20,21</sup>.

Figure 3a shows the mechanical response of the DP-HEA for the coarse-grained (as-homogenized, grain size of  $\sim 45\text{-}\mu\text{m}$ ) and grain-refined (recrystallized, grain size of  $\sim 4.5\text{-}\mu\text{m}$ ) states. To emphasize the substantial improvement in the properties upon grain refinement, the curves for two other single-f.c.c.-phase HEAs ( $\text{Fe}_{37}\text{Mn}_{45}\text{Co}_9\text{Cr}_9$  (ref. 8) and  $\text{Fe}_{20}\text{Mn}_{20}\text{Ni}_{20}\text{Co}_{20}\text{Cr}_{20}$  (ref. 6) (at%)) are also presented.

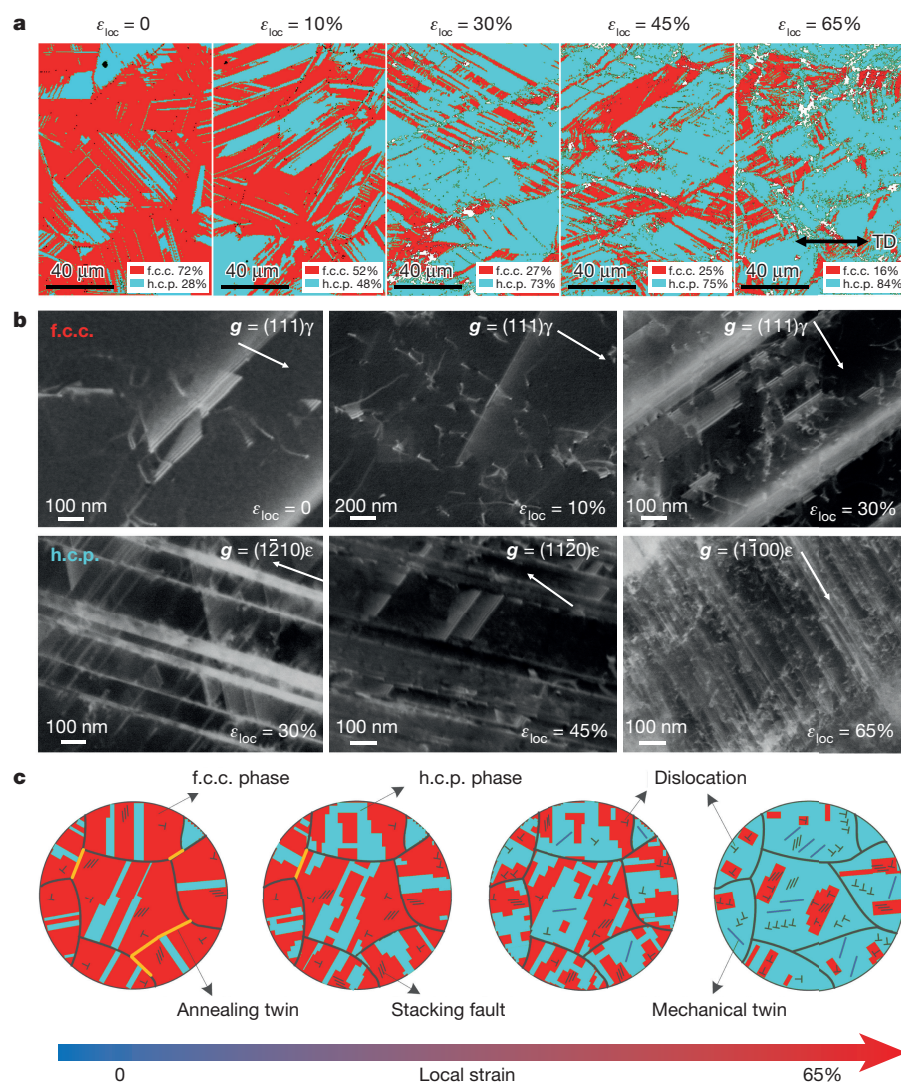
The mechanical response of the TRIP-DP-HEA is striking even before grain refinement. It exhibits vastly higher strength and ductility compared to the single-phase  $\text{Fe}_{37}\text{Mn}_{45}\text{Co}_9\text{Cr}_9$  (at%) HEA. More



**Figure 3 | Mechanical behaviour of the TRIP-DP-HEAs compared to various single-phase HEAs.** Grain sizes are shown in micrometres. **a**, Tensile properties. The tensile-curve data of single-phase  $\text{Fe}_{20}\text{Mn}_{20}\text{Ni}_{20}\text{Co}_{20}\text{Cr}_{20}$  (at%) in ref. 6 is also shown here. The inset shows the increments of (change in) strength and ductility in refs 16 and 17 and this work; ultimate tensile strength and elongation to fracture were used as strength and ductility. Inset labels a, b and c represent the heterogeneous lamella Ti60, Ti80 and Ti100 versus coarse-grained Ti, respectively<sup>16</sup>;

d, e and f represent the high-specific-strength steels I, II and III versus weight-reduced Fe-Al-Mn-C steel, respectively<sup>17</sup>; g and h represent the coarse-grained (grain size  $\sim 45\text{-}\mu\text{m}$ ) and grain refined (grain size  $\sim 4.5\text{-}\mu\text{m}$ ) DP-HEAs versus the single-phase  $\text{Fe}_{37}\text{Mn}_{45}\text{Co}_9\text{Cr}_9$  (at%) HEA, respectively. **b**, Strain-hardening response. The inset shows how the stability of the f.c.c. phase was optimized upon grain refinement to increase the strain-hardening ability; the data points in the inset are means  $\pm$  standard deviation of three tests.





**Figure 4 | Deformation micro-mechanisms in the TRIP-DP-HEA with increasing tensile deformation at room temperature.** **a**, EBSD phase maps revealing the deformation-induced martensitic transformation as a function of deformation.  $\epsilon_{loc}$  is the local strain and TD is the tensile direction. **b**, ECCI analyses showing the evolution of defect substructures in the f.c.c. and h.c.p. phases.  $g$  is the diffraction vector,  $\gamma$  is the f.c.c. phase and  $\epsilon$  is the h.c.p. phase. **c**, Schematic sketches illustrating the sequence of micro-processes in the TRIP-DP-HEA.

importantly, it has a mechanical response almost identical to that of the (grain-refined) single-phase  $\text{Fe}_{20}\text{Mn}_{20}\text{Ni}_{20}\text{Co}_{20}\text{Cr}_{20}$  (at%), the most successful HEA so far<sup>6,22</sup>. On grain refinement, TRIP-DP-HEA notably outperforms the  $\text{Fe}_{20}\text{Mn}_{20}\text{Ni}_{20}\text{Co}_{20}\text{Cr}_{20}$  (at%) HEA. Furthermore, the inset in Fig. 3a demonstrates that this approach in TRIP-DP-HEA has the potential to lead to superior improvements in strength–ductility combinations compared to those obtained in other studies that focus on conventional low-entropy systems<sup>16,17</sup>. Here, the ductility improvements observed in metallic glass matrix composites<sup>23,24</sup> are not shown for comparison, since for these materials the absolute levels of uniform tensile elongation are very low (for example, <3%)<sup>24</sup> even when improved.

We note that the TRIP-DP-HEA reported in this work was designed mainly with the aim of proof of the proposed principle. In our opinion, much more substantial improvements can be achieved following the principles proposed here, if the microstructures and compositions are optimized further. Figure 3b reveals that these improvements correspond to a higher work hardening rate in the DP-HEA  $\text{Fe}_{50}\text{Mn}_{30}\text{Co}_{10}\text{Cr}_{10}$  (at%) than in the single-phase HEAs. This contributes also to an extended uniform deformation process (Extended Data Fig. 2a). There is a notable difference between the strain hardening responses of the coarse-grained and grain-refined  $\text{Fe}_{50}\text{Mn}_{30}\text{Co}_{10}\text{Cr}_{10}$  (at%) HEAs, which is linked to the size dependence of the f.c.c. phase stability (see the inset in Fig. 3b).

These large improvements in the mechanical properties of the TRIP-DP-HEA arise from the underlying plastic accommodation and

hardening processes. By coupling EBSD (Fig. 4a) with electron channeling contrast imaging (ECCI<sup>25</sup>) (Fig. 4b), here we unravel these underlying processes for the case of the coarse-grained TRIP-DP-HEA (Fig. 4c).

We first focus on the f.c.c. phase. EBSD phase maps reveal that the f.c.c. phase is metastable, as desired. It exhibits deformation-stimulated martensitic transformation (f.c.c.  $\rightarrow$  h.c.p.) as a primary deformation mechanism (Fig. 4a). The importance of this mechanism in the observed hardening response can be assessed by comparing the two TRIP-DP-HEAs with different grain sizes: when the stability of the f.c.c. phase is optimized such that martensitic transformation is observed over an extended deformation regime (as for the grain-refined TRIP-DP-HEA; see inset of Fig. 3b), the overall ductility is increased (compare the grain-refined and coarse-grained TRIP-DP-HEAs in Fig. 3a).

The ECCI analysis reveals the evolution of the deformation substructure in the f.c.c. phase (Fig. 4b). Prior to deformation (local strain,  $\epsilon_{loc} = 0$ ), a large number of stacking faults is observed in the TRIP-DP-HEA. Stacking faults present in the f.c.c.  $\gamma$  phase are formed by gliding of Shockley partials of  $1/6\langle 112 \rangle$  Burgers vector<sup>26,27</sup>. These features constitute thin plates of h.c.p. structure (that is, several atomic monolayers of stacking faults). These thin h.c.p. plates have been shown to act as the nuclei of the  $\epsilon$  martensite phase<sup>27</sup>, which forms through the overlapping of stacking faults. The observed stacking faults in the undeformed HEA are initial faults that did not sufficiently coalesce to form the thermally induced h.c.p.  $\epsilon$  phase but are likely to act as phase-formation nuclei when subjected to externally applied mechanical loads.



At early stages of deformation, mechanically induced transformation from the f.c.c.  $\gamma$  phase to the h.c.p.  $\epsilon$  phase acts as the primary deformation mechanism (Fig. 4a). Since the stacking faults act as nuclei for the formation of the h.c.p.  $\epsilon$  phase, a large number of stacking faults are required in the f.c.c. phase to realize the transformation from f.c.c. to h.c.p. phase at this stage. This is well documented by ECCI, together with an increase in the dislocation density (see the  $\epsilon_{\text{loc}} = 10\%$  and  $\epsilon_{\text{loc}} = 30\%$  states in Fig. 4b). Thus, dislocation plasticity and martensitic transformation plasticity are both activated at similar deformation levels. The increased phase boundary density due to transformation creates additional obstacles of dislocation slip, thereby contributing to the strain hardening. With increasing strain, transformation from the f.c.c.  $\gamma$  to the h.c.p.  $\epsilon$  phase continues to be the dominant deformation mechanism, yet dislocation activity in the f.c.c.  $\gamma$  phase becomes more important. At 65% local strain (corresponding to the post-necking state, see Extended Data Fig. 2b), only  $\sim 16\%$  of the f.c.c. phase is retained (Fig. 4a).

We next focus on the h.c.p. phase. The TRIP-DP-HEA was water-quenched after homogenization at 1,200 °C, so the starting h.c.p.  $\epsilon$  phase is thermally induced by the martensitic transformation. Neither the thermally induced  $\epsilon$  regions nor the mechanically induced h.c.p. regions show notable deformation-induced features at low strain levels. Gradually, deformation-induced twinning was observed in the h.c.p.  $\epsilon$  phase as an important deformation mechanism (Fig. 4b). Thus, as the local strain increases to 30% and then to 45%, an increase in the density of both mechanical nano-twins and stacking faults is observed in the h.c.p. phase (Fig. 4b). The phenomenon of twinning in the deformation-induced h.c.p. martensite has also been observed in other types of alloys<sup>28</sup>. This mechanism contributes profoundly to strain hardening through the dynamic Hall–Petch effect. This means that the interface density, through the continuously formed twins, increases constantly<sup>29</sup>. Further increase of the deformation leads to the presence of a high density of dislocations (Fig. 4b). Thus, the h.c.p.  $\epsilon$  phase plays an important part in plastic accommodation and hardening at later stages of deformation via multiple deformation mechanisms (that is, dislocation slip, twinning and the formation of stacking faults).

These deformation micro-mechanisms (Fig. 4) and the impressive mechanical response (Fig. 3) confirm the success of this method of simultaneously achieving greatly improved strength (from massive solid solution strengthening and the increased interface density) and ductility (from dislocation-plasticity and transformation-induced hardening). The synergic deformation of the two phases leads to a highly beneficial dynamic strain–stress partitioning effect<sup>30</sup>; with a decreased likelihood of damage nucleation owing to their elastic compliance. Such damage resistance is absent in most dual-phase alloys with high mechanical contrast across their hetero-interfaces<sup>12</sup>.

Our effort to combine the best characteristics of steels and HEAs has led to the design of a new class of transformation-induced plasticity-assisted, dual-phase HEA. The originally proposed HEA concept has motivated enormous efforts to design new alloys, but few of the resulting alloys have shown properties that justify the increased alloying content, in contrast to the alloy presented here, which exhibits excellent strength–ductility combinations. We emphasize that this alloy design strategy is opposite in approach to that generally used in HEAs design: rather than focusing on phase stabilization and single-phase formation, we propose that phase metastability, and ductile multi-phase configurations should be important future research goals in this field.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 2 November 2015; accepted 24 March 2016.

Published online 18 May 2016.

1. Ritchie, R. O. The conflicts between strength and toughness. *Nature Mater.* **10**, 817–822 (2011).
2. Wei, Y. *et al.* Evading the strength–ductility trade-off dilemma in steel through gradient hierarchical nanotwins. *Nature Commun.* **5**, 3580 (2014).

3. Yeh, J. W. *et al.* Nanostructured high-entropy alloys with multiple principal elements: novel alloy design concepts and outcomes. *Adv. Eng. Mater.* **6**, 299–303 (2004).
4. Zhang, Y. *et al.* Microstructures and properties of high-entropy alloys. *Prog. Mater. Sci.* **61**, 1–93 (2014).
5. Gludovatz, B. *et al.* Exceptional damage-tolerance of a medium-entropy alloy CrCoNi at cryogenic temperatures. *Nature Commun.* **7**, 10602 (2016).
6. Gludovatz, B. *et al.* A fracture-resistant high-entropy alloy for cryogenic applications. *Science* **345**, 1153–1158 (2014).
7. Yao, M. J., Pradeep, K. G., Tسان, C. C. & Raabe, D. A novel, single phase, non-equiatomic FeMnNiCoCr high-entropy alloy with exceptional phase stability and tensile ductility. *Scr. Mater.* **72–73**, 5–8 (2014).
8. Tسان, C. C. *et al.* Composition dependence of phase stability, deformation mechanisms, and mechanical properties of the CoCrFeMnNi high-entropy alloy system. *JOM* **66**, 1993–2001 (2014).
9. Pradeep, K. G. *et al.* Non-equiatomic high entropy alloys: approach towards rapid alloy screening and property-oriented design. *Mater. Sci. Eng. A* **648**, 183–192 (2015).
10. Deng, Y. *et al.* Design of a twinning-induced plasticity high entropy alloy. *Acta Mater.* **94**, 124–133 (2015).
11. Wang, Y. P., Li, B. S. & Fu, H. Z. Solid solution or intermetallics in a high-entropy alloy. *Adv. Eng. Mater.* **11**, 641–644 (2009).
12. Tسان, C. C. *et al.* An overview of dual-phase steels: advances in microstructure-oriented processing and micromechanically guided design. *Annu. Rev. Mater. Res.* **45**, 391–431 (2015).
13. Herrera, C., Ponge, D. & Raabe, D. Design of a novel Mn-based 1 GPa duplex stainless TRIP steel with 60% ductility by a reduction of austenite stability. *Acta Mater.* **59**, 4653–4664 (2011).
14. Hadfield, R. A. Hadfield's manganese steel. *Science* **12**, 284–286 (1888).
15. Grässel, O., Krüger, L., Frommeyer, G. & Meyer, L. W. High strength Fe–Mn–(Al, Si) TRIP/TWIP steels development–properties–application. *Int. J. Plast.* **16**, 1391–1409 (2000).
16. Wu, X. *et al.* Heterogeneous lamella structure unites ultrafine-grain strength with coarse-grain ductility. *Proc. Natl Acad. Sci. USA* **112**, 14501–14505 (2015).
17. Kim, S.-H., Kim, H. & Kim, N. J. Brittle intermetallic compound makes ultrastrong low-density steel with large ductility. *Nature* **518**, 77–79 (2015).
18. Pierce, D. T. *et al.* The influence of manganese content on the stacking fault and austenite/ $\epsilon$ -martensite interfacial energies in Fe–Mn–(Al–Si) steels investigated by experiment and theory. *Acta Mater.* **68**, 238–253 (2014).
19. Mandal, S., Pradeep, K. G., Zaefferer, S. & Raabe, D. A novel approach to measure grain boundary segregation in bulk polycrystalline materials in dependence of the boundaries' five rotational degrees of freedom. *Scr. Mater.* **81**, 16–19 (2014).
20. Dmitrieva, O. *et al.* Chemical gradients across phase boundaries between martensite and austenite in steel studied by atom probe tomography and simulation. *Acta Mater.* **59**, 364–374 (2011).
21. Raabe, D. *et al.* Segregation engineering enables nanoscale martensite to austenite phase transformation at grain boundaries: a pathway to ductile martensite. *Acta Mater.* **61**, 6132–6152 (2013).
22. Otto, F. *et al.* The influences of temperature and microstructure on the tensile properties of a CoCrFeMnNi high-entropy alloy. *Acta Mater.* **61**, 5743–5755 (2013).
23. Hays, C., Kim, C. & Johnson, W. L. Microstructure controlled shear band pattern formation and enhanced plasticity of bulk metallic glasses containing in situ formed ductile phase dendrite dispersions. *Phys. Rev. Lett.* **84**, 2901 (2000).
24. Hofmann, D. C. *et al.* Designing metallic glass matrix composites with high toughness and tensile ductility. *Nature* **451**, 1085–1089 (2008).
25. Zaefferer, S. & Elhami, N.-N. Theory and application of electron channelling contrast imaging under controlled diffraction conditions. *Acta Mater.* **75**, 20–50 (2014).
26. Yakubtsov, I. A., Ariapour, A. & Perovic, D. D. Effect of nitrogen on stacking fault energy of f.c.c. iron-based alloys. *Acta Mater.* **47**, 1271–1279 (1999).
27. Brooks, J. W., Loretto, M. H. & Smallman, R. E. Direct observations of martensite nuclei in stainless steel. *Acta Metall.* **27**, 1839–1847 (1979).
28. Kim, C. P., Oh, Y. S., Lee, S. & Kim, N. J. Realization of high tensile ductility in a bulk metallic glass composite by the utilization of deformation-induced martensitic transformation. *Scr. Mater.* **65**, 304–307 (2011).
29. Lu, K., Lu, L. & Suresh, S. Strengthening materials by engineering coherent internal boundaries at the nanoscale. *Science* **324**, 349–352 (2009).
30. Wang, M. M., Tسان, C. C., Ponge, D., Dippel, A. C. & Raabe, D. Nanolaminate transformation-induced plasticity–twinning-induced plasticity steel with dynamic strain partitioning and enhanced damage resistance. *Acta Mater.* **85**, 216–228 (2015).

**Acknowledgements** This work is financially supported by the European Research Council under the EU's 7th Framework Programme (FP7/2007-2013)/ERC grant agreement 290998. The contributions of H. Springer, S. Zaefferer, M. Nellessen, M. Adamek and F. Schlüter are also gratefully acknowledged.

**Author Contributions** C.C.T. and D.R. designed the research; Z.L. was the lead experimental scientist of the study; K.G.P. and Y.D. performed some of the alloy design experiments; and Z.L. and C.C.T. wrote the paper. All authors discussed the results and commented on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.R. ([d.raabe@mpie.de](mailto:d.raabe@mpie.de)), C.C.T. ([tsan@mit.edu](mailto:tsan@mit.edu)) or Z.L. ([zhiming.li@mpie.de](mailto:zhiming.li@mpie.de)).

## METHODS

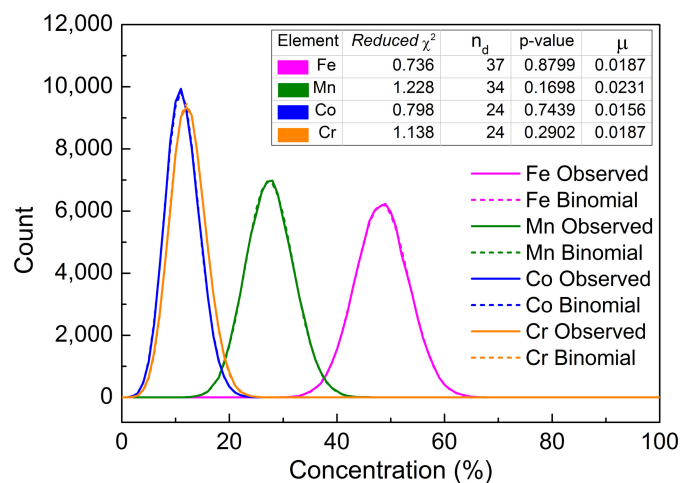
During the processing of the HEAs presented in this work, the ingot was first cast in a vacuum induction furnace using pure metals (purity higher than 99.8 wt%) with predetermined compositions. The as-cast ingot with dimensions of  $10 \times 50 \times 150 \text{ mm}^3$  was subsequently hot-rolled at  $900^\circ\text{C}$  with a rolling reduction ratio of 50% (thickness changed from 10 mm to 5 mm). After hot-rolling, the alloy was homogenized at  $1,200^\circ\text{C}$  for 2 h in an Ar atmosphere followed by water-quenching. For the  $\text{Fe}_{50}\text{Mn}_{30}\text{Co}_{10}\text{Cr}_{10}$  (at%) DP-HEA, further grain refinement was achieved through cold-rolling with a reduction ratio of 60% and subsequent recrystallization annealing at  $900^\circ\text{C}$  in an Ar atmosphere for 3 min followed by water-quenching. The bulk chemical compositions of all the studied alloys were measured by wet-chemical analysis (Extended Data Table 1).

The microstructures of the alloys were analysed using multiple techniques. EBSD measurements were performed using a Zeiss-Crossbeam XB 1540 focused ion beam scanning electron microscope (SEM) with a Hikari camera and the TSL OIM data-collection software (<http://www.edax.com/Products/EBSD/OIM-Data-Collection-EBSD-SEM.aspx>). Back-scattered electron imaging and ECCI<sup>25</sup> analyses were carried out using a Zeiss-Merlin instrument. The chemical uniformity

was investigated using energy-dispersive X-ray spectroscopy at the microscopic scale, and APT (LEAP 3000X HR, Cameca Inc.) at the atomic scale. The APT tips were produced using a focused ion beam (FEI Helios Nanolab 600i) from regions including phase and grain boundaries revealed by a prior EBSD scan.

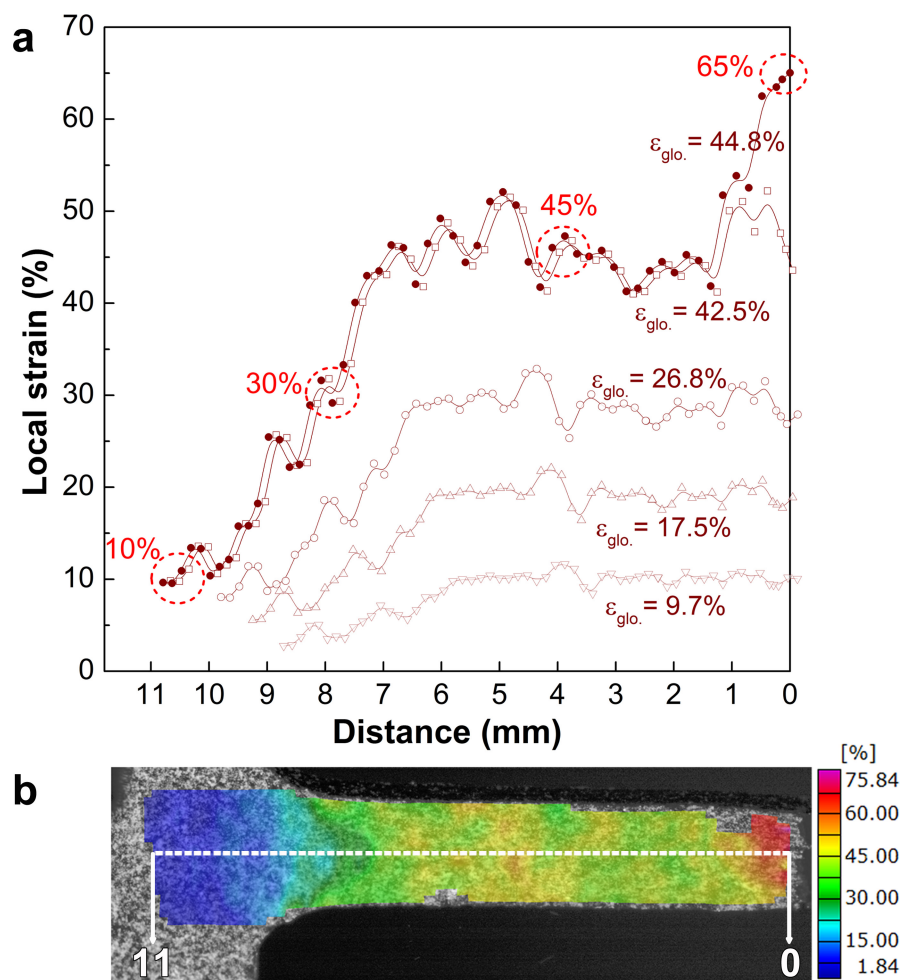
Flat specimens for tensile testing, with a thickness of 1 mm, were sectioned from the homogenized and water-quenched alloy by electrical discharge machining. The gauge length and width of the tensile specimens were 10 mm and 2.5 mm, respectively. Uniaxial tensile tests were carried out at ambient temperature using a Kammrath & Weiss tensile stage at the strain rate of  $1 \times 10^{-3} \text{ s}^{-1}$ . Five samples for each material were tensile-tested to confirm reproducibility. The local strain evolution during tensile test was determined by digital image correlation using the Aramis system (GOM GmbH, <http://www.gom.com/metrology-systems/system-overview/aramis.html>).

The deformation mechanisms in the DP-HEAs were investigated by EBSD and ECCI at different regions of the fractured tensile sample with different local strain levels. All of the sample regions analysed by ECCI were first measured by EBSD to obtain the specific orientation information corresponding to each region.



**Extended Data Figure 1 | Statistical binomial frequency distribution analysis results for the APT tip.** The statistical analysis shows that the tip has an overall composition of  $\text{Fe}_{48.6}\text{Mn}_{27.6}\text{Co}_{11.3}\text{Cr}_{12.3}$  (at%). The binomial curves obtained from the experiments match the curves corresponding to a total random distribution. The quality of the fit was quantified using several parameters, as listed in the key.  $n_d$  is the number of degrees of freedom for a given ion. The values of the normalized homogenization parameter  $\mu$  for all four elements are close to 0, confirming the random distribution of elements in the DP-HEA.





**Extended Data Figure 2 | Strain distribution within the DP-HEA sample upon room-temperature deformation. a,** Evolution of local strain with increasing the global strain ( $\epsilon_{\text{glo.}}$ ), indicating an extended uniform deformation process. The red dotted circles in **a** indicate the local strain values corresponding to various positions in the fractured tensile sample shown in **b**; four positions with local strains of 10%, 30%, 45%

and 65% were highlighted by percentages in red and the corresponding microstructures are shown in Fig. 4. **b,** Digital image correlation strain map shows the local strain distribution of the tensile sample following fracture. 0 to 11 in **b** refers to the distance of the sample position from the fracture surface, corresponding to the distance values shown in **a**.

Extended Data Table 1 | Chemical composition of the studied alloys in atomic per cent according to wet-chemical analysis

Alloy	Fe	Mn	Co	Cr
45%Mn	37.10	45.00	8.83	9.07
40%Mn	39.82	41.60	9.17	9.40
35%Mn	44.44	35.09	10.15	10.31
30%Mn	47.31	31.38	10.54	10.77

# Strongly correlated perovskite fuel cells

You Zhou<sup>1</sup>, Xiaofei Guan<sup>1</sup>, Hua Zhou<sup>2</sup>, Koushik Ramadoss<sup>1</sup>, Suhare Adam<sup>1</sup>, Huajun Liu<sup>3</sup>, Sungsik Lee<sup>2</sup>, Jian Shi<sup>1,4</sup>, Masaru Tsuchiya<sup>5</sup>, Dillon D. Fong<sup>3</sup> & Shriram Ramanathan<sup>1,6</sup>

**Fuel cells convert chemical energy directly into electrical energy with high efficiencies and environmental benefits, as compared with traditional heat engines<sup>1–4</sup>. Yttria-stabilized zirconia is perhaps the material with the most potential as an electrolyte in solid oxide fuel cells (SOFCs), owing to its stability and near-unity ionic transference number<sup>5</sup>. Although there exist materials with superior ionic conductivity, they are often limited by their ability to suppress electronic leakage when exposed to the reducing environment at the fuel interface. Such electronic leakage reduces fuel cell power output and the associated chemo-mechanical stresses can also lead to catastrophic fracture of electrolyte membranes<sup>6</sup>. Here we depart from traditional electrolyte design that relies on cation substitution to sustain ionic conduction. Instead, we use a perovskite nickelate as an electrolyte with high initial ionic and electronic conductivity. Since many such oxides are also correlated electron systems, we can suppress the electronic conduction through a filling-controlled Mott transition induced by spontaneous hydrogen incorporation. Using such a nickelate as the electrolyte in free-standing membrane geometry, we demonstrate a low-temperature micro-fabricated SOFC with high performance. The ionic conductivity of the nickelate perovskite is comparable to the best-performing solid electrolytes in the same temperature range, with a very low activation energy. The results present a design strategy for high-performance materials exhibiting emergent properties arising from strong electron correlations.**

SmNiO<sub>3</sub> (SNO) belongs to a series of rare-earth nickelates (RNiO<sub>3</sub> or RNO) with the perovskite structure (ABO<sub>3</sub>), which exhibits linked corner-shared BO<sub>6</sub> octahedra (Fig. 1a)<sup>7</sup>. In perovskite oxides, protons can form ionic defects (OH<sub>O</sub> in Kröger–Vink notation) by bonding with oxygen<sup>8</sup>, and diffuse through a Grotthuss mechanism that involves the fast rotational diffusion of the protonic defects and the rate-limiting proton transfer to the neighbouring oxygen ions<sup>8,9</sup>. The transition states of the proton rotation and proton transfer require local lattice distortions such as elongation and bending of the B–O bond, respectively<sup>10,11</sup>. The schematic of proton incorporation and diffusion processes for a cubic perovskite is shown in Fig. 1b with the following processes: (i) proton incorporation, (ii) rotational diffusion, (iii) transfer to neighbouring oxygen, (iv) bending and (v) elongation of the B–O bond. In SNO, proton incorporation and diffusion happen in a similar way, albeit with several different characteristics, as will be discussed in more detail later (Fig. 1c).

In the low-temperature fuel cell operation range (300–500 °C), stoichiometric SNO shows metallic conductivity with an electrical resistivity of ~1 mΩ cm, which is detrimental to electrolyte applications. The high electronic conductivity is due to single electron occupancy on the fourfold degenerate *e<sub>g</sub>* manifold (including spin) on Ni<sup>3+</sup>, as shown in Fig. 1d (in the ionic limit; the covalent limit cases are shown in Extended Data Fig. 1a and b), where carriers can migrate without overcoming the on-site Coulomb repulsion. When electrons are doped

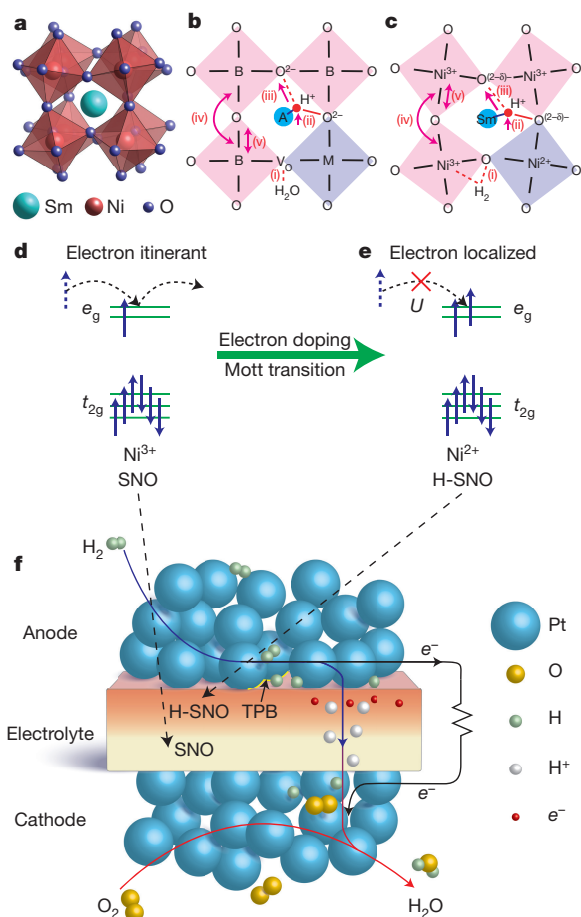
into SNO via hydrogenation and the valence of nickel is reduced to Ni<sup>2+</sup> (with overall reaction  $\text{Ni}^{3+} + \text{O}_\text{O}^\times + \frac{1}{2}\text{H}_2 \leftrightarrow \text{Ni}^{2+} + \text{OH}_\text{O}^\cdot$ ), however, electronic transport through the *e<sub>g</sub>* manifold will be suppressed by the Hubbard intra-orbital electron–electron Coulomb interaction *U* (Fig. 1e). Such filling-controlled Mott transitions enable the application of hydrogenated SNO as an electrolyte, owing to its wide electronic bandgap<sup>12</sup>, which is close to the Ni intra-orbital Coulomb repulsion and large enough to suppress electronic conductivity<sup>13</sup>. Spontaneous incorporation of protons into SNO upon hydrogen exposure without any electrical bias at low temperatures can be seen in Extended Data Fig. 2. This is unlike typical perovskite proton conductors such as yttrium-doped BaCeO<sub>3</sub> and BaZrO<sub>3</sub>, where subvalent cations are needed as substitutional acceptors to facilitate the hydrogen-incorporation process (Fig. 1b). Therefore the concentration of protons in SNO may not be limited by the oxygen vacancy concentration, as commonly noted in acceptor-doped electrolytes. The electronic transport mechanism in H-SNO is characterized by the Efros–Shklovskii variable range hopping mechanism, in which small polarons form because of strong electron–lattice coupling in the presence of a Coulomb gap (Extended Data Fig. 1c–e).

Figure 1f illustrates how this collective quantum mechanical effect enables the electrolyte design. Initially no power output is extracted from the SNO-electrolyte fuel cell because of the high electronic conductivity in pristine SNO. When the hydrogen fuel is introduced at the anode (catalytic Pt or Pd), hydrogen molecules dissociate into protons and donate electrons to Ni(III) in SNO at the triple phase boundaries. The hydrogenation process creates an electrically insulating H-SNO on the anode side. Once this insulating layer is formed, as long as hydrogen fuel is supplied, protons can continue to diffuse under the chemical potential gradient, while the electron transport through H-SNO directly to the cathode is strongly suppressed by carrier localization. As a result, electrons are forced to pass through the external circuit and generate electrical power.

The time evolution of the open-circuit voltage (OCV) in a micro-fabricated SOFC with a free-standing SNO membrane (see Extended Data Figs 3 and 4 for the device structure and fabrication) as the electrolyte verifies the above mechanism (Extended Data Fig. 5a). Initially there is no OCV as the cell is electrically shorted by pristine SNO. The OCV increases under continuous hydrogen flow after the temperature becomes stabilized, as the H-SNO phase forms on the anode side, and reaches a stable output when the stationary state is reached. The current–voltage characteristics of the micro-fabricated SOFCs (Fig. 2a) exhibit typical activation polarization, ohmic loss and concentration polarization behaviour, and the power output reaches a maximum value of 225 mW cm<sup>−2</sup> at 500 °C, which is comparable to the best-performing proton conducting fuel cells (ref. 14 and references therein). The highest OCV achieved (1.03 V) is close to the Nernst potential (~1.07 V), showing that the ionic transference number is close to unity, with the electronic conduction

<sup>1</sup>John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA. <sup>2</sup>X-ray Science Division, Advanced Photon Source, Argonne National Laboratory, Argonne, Illinois 60439, USA. <sup>3</sup>Materials Science Division, Argonne National Laboratory, Argonne, Illinois 60439, USA. <sup>4</sup>Department of Materials Science and Engineering, Rensselaer Polytechnic Institute, Troy, New York 12180, USA. <sup>5</sup>SiEnergy Systems, Cambridge, Massachusetts 02140, USA. <sup>6</sup>School of Materials Engineering, Purdue University, West Lafayette, Indiana 47907, USA.





**Figure 1 | Solid electrolyte design principle based on the emergent phase arising from strong correlations.** **a**, The distorted perovskite structure of the SNO crystal. **b**, **c**, Proton incorporation and conduction mechanisms in a conventional solid-state electrolyte (A, B and M are metal cations and O is the oxygen anion) (**b**) and the proposed new electrolyte (**c**). (i) Proton incorporation. (ii) and (iii), Proton transport by rotational diffusion within an octahedron (ii) and transfer to a neighbouring oxygen ion facilitated by the hydrogen bond (dashed red line) (iii). (iv) and (v), The bending (iv) and the stretching (v) of the metal–oxygen bond promote processes (iii) and (ii), respectively. In conventional electrolytes, substitutional sub-valent cations, M, are needed to facilitate the hydrogen incorporation. In SNO, proton incorporation can happen spontaneously. The ligand holes in SNO reduce the effective charge of the oxygen ions (only two of them are explicitly shown). **d**, **e**, The electronic configuration of Ni 3d orbitals for the pristine (**d**) and the electron-doped (**e**) SNO in the ionic limit. Electronic transport is suppressed by the on-site electron–electron correlation  $U$  upon electron doping (**e**). **f**, A schematic of a SNO-electrolyte SOFC and its operation mechanism. Spontaneous hydrogen incorporation creates a strongly correlated insulating layer and suppresses the electronic current. TPB, triple phase boundary.

almost completely suppressed. The deviation of OCV from the Nernst potential in general could be related to gas leakage and residual electronic conductivity. The ionic transference number of H-SNO at 500 °C is estimated to be 0.96 using the standard electromotive force method. Because there is a small yet finite current inside the fuel cells through the electrolyte under the OCV condition, the electrode polarization loss (in addition to ohmic loss) may contribute to the deviation of the measured OCV from the ideal value. Therefore, a method that considers the polarization loss may also be used to evaluate the ionic transference number<sup>15,16</sup> (Extended Data Fig. 5b). Increasing the electrolyte thickness typically enhances the measured OCV, possibly owing to the reduced possibilities of pinholes in the membrane and a decrease in the relative ratio between electrode polarization and electrolyte resistance (see Extended

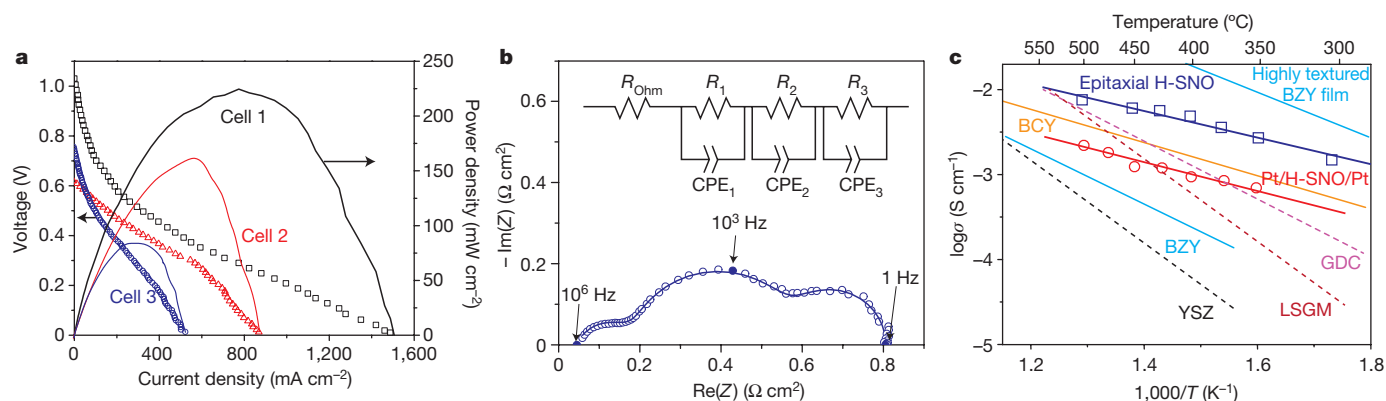
Data Fig. 6a and additional discussions in the Supplementary Information).

H-SNO fuel cells with dense Pd anodes also produce power output, indicating that protons rather than oxygen ions are the dominant mobile ion species in the material (Extended Data Fig. 6b). In addition, these fuel cells can also work under pure H<sub>2</sub> and are stable over tens of hours (Extended Data Figs 6b and 7). The Nyquist plot of the cell under OCV conditions measured at 500 °C is shown in Fig. 2b. The extrapolated area specific resistance of H-SNO, one of the key performance metrics, is remarkably low: 0.045 Ω cm<sup>2</sup> at 500 °C, which is less than one-third of the general target value (0.15 Ω cm<sup>2</sup>) for the area specific resistance of an oxide electrolyte<sup>17</sup>. The Nyquist plot can be modelled by an equivalent circuit with an ohmic resistor,  $R_{\text{ohm}}$ , and serial elements each consisting of a resistor,  $R_i$  ( $i = 1, 2, 3$ ), and a constant phase element,  $\text{CPE}_i$  ( $i = 1, 2, 3$ ), as summarized in Supplementary Table 1. The three semicircles in the Nyquist plot originate from the electric double-layer capacitance at the anode and cathode, and the pseudocapacitance related to hydrogen incorporation in SNO (see the footnote to Supplementary Table 1).

Figure 2c shows the ionic conductivity (calculated by the electromotive force method) of H-SNO measured from free-standing Pt/H-SNO/Pt micro-fabricated SOFCs and H-SNO epitaxial films on LaAlO<sub>3</sub> (LAO) (001) (indexed in pseudocubic notation), compared with several other best-performing oxygen-ion-conducting and proton-conducting electrolyte materials<sup>17–21</sup>. SNO has a high ionic conductivity with low activation energy ( $\sim 0.3$  eV, similar to solid acid protonic conductors<sup>22</sup>), making it especially suitable for low-temperature SOFC applications<sup>3</sup>. The difference in the ionic conductivity measured from epitaxial thin films and membranes could be related to contributions from grain boundaries<sup>20,23</sup>. Grain boundaries may not only decrease proton mobility by scattering and trapping, but may also reduce the proton concentration proximal to the boundaries by creating space charge layers. Therefore the total ionic resistance of polycrystalline samples can be larger than that of the epitaxial films.

Several factors may collectively lead to the high ionic conductivity with low activation energy in SNO. First, it has been found that in RNO, Ni forms a covalent bond with O in a mixed electron configuration of  $3d^7$  and  $3d^8\bar{L}$  (where  $\bar{L}$  denotes a ligand hole on O 2p) (ref. 24). The covalence reduces the effective charge on oxygen and therefore the bonding strength between the oxygen ion and the proton, which lowers the proton transfer activation energy (Fig. 1c). Additionally, the proton transport barrier in perovskites with a tetravalent B-site ( $A(\text{II})\text{--}B(\text{IV})$ ) is in general much smaller than the ones with a pentavalent B-site ( $A(\text{I})\text{--}B(\text{V})$ )<sup>25</sup>. It has been suggested that  $A(\text{III})\text{--}B(\text{III})$  perovskites may have even higher ionic conductivity<sup>25</sup>. This may be explained by the weaker repulsion between B-site ions and protons in  $A(\text{III})\text{--}B(\text{III})$  perovskites, which reduces the energy of the proton in its transition state. Finally, as the transition states of the proton rotation and proton transfer require local lattice distortions such as elongation and bending of the B–O bond, respectively<sup>10,11</sup>, the relative low energy of the Ni–O bending and stretching modes in SNO ( $\sim 35$  meV and 75 meV, respectively<sup>26</sup>) can also contribute to lowering the proton transport barrier.

To confirm the electron localization mechanism during fuel cell operation and to reveal the underlying reasons for the high ionic conductivity, both chemical and structural characterizations of the SNO hydrogenation process were performed. *Ex situ* X-ray absorption near-edge spectroscopy (XANES) measurements of the nickel K-edge from a pristine and a hydrogenated SNO sample are shown in Fig. 3a, as well as that from a reference nickel metal sample used for energy calibration. Several features are present in the spectra of SNO and H-SNO. The pre-edge feature, A, originates from the dipolar transition between Ni 1s and Ni 3d–O 2p hybridized  $3d^8\bar{L}$ , and points to the covalent nature of the Ni–O bond<sup>27</sup>. Features B, D and E are derived from the first oxygen coordination shell, while C and C' originate from the second shell of the rare-earth ions<sup>27</sup>.



**Figure 2 | Performance of the emergent-phase electrolyte in fuel cells.** **a**, Typical current density–voltage characteristics and the power densities of Pt/H-SNO/Pt micro-fabricated SOFCs measured at 500 °C with 3% humidified 5% H<sub>2</sub>–95% Ar as fuel and laboratory air as oxidant. The electrolyte thickness is 1.5 μm for cell 1, and 1 μm for cells 2 and 3. **b**, A Nyquist plot measured under OCV conditions at 500 °C for a Pt/SNO/Pt cell (solid line shows the fitted curve).  $Z$  is the complex impedance measured from the fuel cell. The Nyquist plot can be modelled by an equivalent circuit (inset) with an ohmic resistor,  $R_{\text{ohm}}$ , and serial

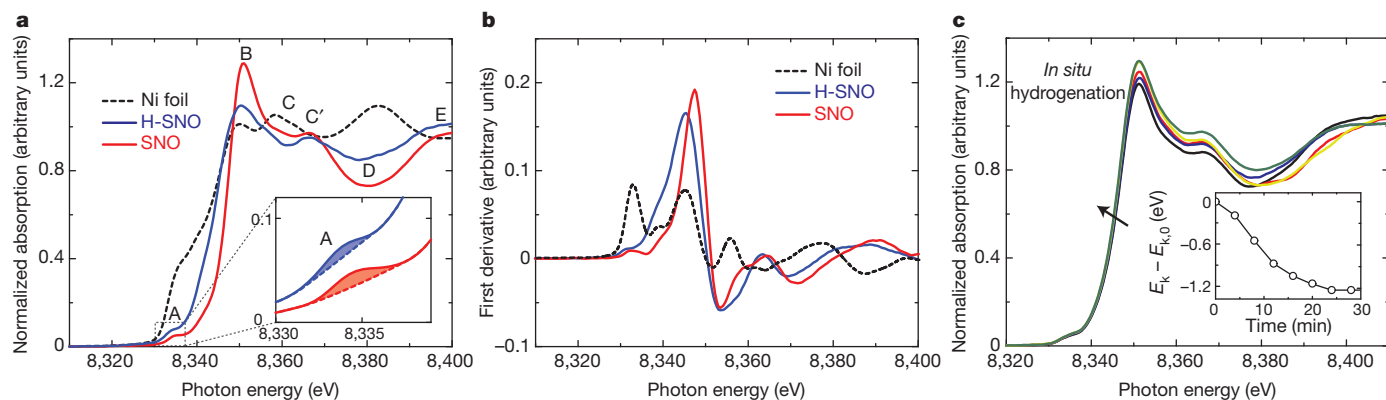
elements each consisting of a resistor,  $R_i$  ( $i = 1, 2, 3$ ), and a constant phase element,  $\text{CPE}_i$  ( $i = 1, 2, 3$ ). **c**, The ionic conductivity of H-SNO compared to the best-performing oxygen-ion-conducting electrolytes (dashed lines) and proton conductors (solid lines). The oxygen-ion-conducting electrolytes are: stabilized zirconia (YSZ, (ZrO<sub>2</sub>)<sub>0.9</sub>(Y<sub>2</sub>O<sub>3</sub>)<sub>0.1</sub>) (ref. 17), La<sub>0.8</sub>Sr<sub>0.2</sub>Ga<sub>0.8</sub>Mg<sub>0.2</sub>O<sub>3</sub> (LSGM) (ref. 18) and doped ceria (GDC, Ce<sub>0.8</sub>Gd<sub>0.2</sub>O<sub>1.9-δ</sub>) (ref. 19). The proton conductors are BaZr<sub>0.8</sub>Y<sub>0.2</sub>O<sub>3-δ</sub> (BZY, in the form of both sintered pellets and highly textured films) (ref. 20) and BaCe<sub>0.8-x</sub>Zr<sub>x</sub>Y<sub>0.2</sub>O<sub>3-δ</sub> (BCY, 0 <  $x$  < 0.8) (ref. 21).

A substantial shift of the absorption edge to a lower energy is observed upon hydrogenation. The energies of the absorption edge and other features are consistent with those of Ni(III) in RNO<sup>28</sup>. From the inflection point in the first derivative of the absorption (Fig. 3b), the chemical shift from SNO to H-SNO is determined to be ~2.0 eV. A linear relation between the absorption edge and the formal valence state has been previously noted with a slope from ~1.5 eV per electron to ~2.8 eV per electron<sup>28–30</sup>. The absorption inflection point, however, depends on both the Ni valence state and the atomic arrangement around Ni. In this study, the valence change is the primary factor leading to the absorption edge shift, and the Ni valence state change is close to –1, suggesting high proton concentration in H-SNO without introducing impurity dopants. The change of Ni valence state verifies that hydrogen exists as protons in H-SNO, because it is more favourable for Ni to accept an electron from hydrogen rather than from O<sup>2-</sup> and Sm<sup>3+</sup> ions when changing from SmNi(III)O<sub>3</sub> and H-SmNi(II)O<sub>3</sub>. The angle-resolved XANES spectra show that the proton incorporation not only happens at the surface but also through the thickness of the films (Extended Data Fig. 8). In addition, the decrease in the white line intensity suggests an overall decrease in the hole density on Ni after hydrogenation.

The intensity of the pre-edge feature A, offset with respect to feature B by ~14.4 eV, represents the density of the 3d<sup>8</sup>L state in the ground state and decreases upon hydrogenation (inset of Fig. 3a), which shows that the doped electrons partially fill the ligand holes. Following the analysis in ref. 27, we find the pristine ground state to be ~0.5|3d<sup>7</sup>⟩ + 0.5|3d<sup>8</sup>L⟩, and estimate that the concentration of ligand holes decreases by ~50% after the hydrogenation. This verifies that ligand holes are present on oxygen ions in both SNO and H-SNO, which helps to reduce the proton transfer activation energy.

Figure 3c shows the representative *in situ* XANES spectra during the hydrogenation process. The chemical shift is smaller than those of the *ex situ* experiments owing to the lower operation temperature limited by the apparatus. The dynamic change in the absorption edge (inset of Fig. 3c) shows that the average valence state reaches equilibrium in ~30 min at 200 °C.

Synchrotron X-ray diffraction studies (Extended Data Figs 9 and 10) suggest that the SNO lattice expands during hydrogenation, which may lead to a change in the relative rate of inter- and intra-octahedron proton transfer<sup>9</sup> and the lattice open volume, and therefore modify the long-range proton transport properties.



**Figure 3 | Ex situ and in situ XANES characterizations of the phase evolution.** **a**, *Ex situ* normalized Ni K-edge XANES spectra of SNO, H-SNO and the nickel metal reference, with zoomed view of the pre-edge feature 'A' (inset). The other features (B, C, C', D and E) are derived from the first oxygen coordination shell and the second shell of the rare-earth

ions. **b**, First derivative of the normalized absorption. **c**, *In situ* XANES spectra of the SNO hydrogenation process performed at 200 °C. The arrow indicates the direction of time evolution. The dynamics of the shift in the energy of the absorption edge  $E_k$  is shown in the inset (where  $E_{k,0}$  is the absorption edge energy of pristine SNO).

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 13 September 2015; accepted 10 March 2016.**

**Published online 16 May 2016.**

- O'Hayre, R. P., Cha, S.-W., Colella, W. & Prinz, F. B. *Fuel Cell Fundamentals* (John Wiley & Sons, 2006).
- Singhal, S. Advances in solid oxide fuel cell technology. *Solid State Ion.* **135**, 305–313 (2000).
- Wachsman, E. D. & Lee, K. T. Lowering the temperature of solid oxide fuel cells. *Science* **334**, 935–939 (2011).
- Shao, Z. *et al.* A thermally self-sustained micro solid-oxide fuel-cell stack with high power density. *Nature* **435**, 795–798 (2005).
- Haile, S. M. Fuel cell materials and components. *Acta Mater.* **51**, 5981–6000 (2003).
- Kerman, K., Lai, B.-K. & Ramanathan, S. Free standing oxide alloy electrolytes for low temperature thin film solid oxide fuel cells. *J. Power Sources* **202**, 120–125 (2012).
- Catalan, G. Progress in perovskite nickelate research. *Phase Transit.* **81**, 729–749 (2008).
- Kreuer, K.-D. Proton conductivity: materials and applications. *Chem. Mater.* **8**, 610–641 (1996).
- Münch, W., Kreuer, K.-D., Seifert, G. & Maier, J. Proton diffusion in perovskites: comparison between BaCeO<sub>3</sub>, BaZrO<sub>3</sub>, SrTiO<sub>3</sub>, and CaTiO<sub>3</sub> using quantum molecular dynamics. *Solid State Ion.* **136/137**, 183–189 (2000).
- Münch, W., Seifert, G., Kreuer, K. D. & Maier, J. A quantum molecular dynamics study of proton conduction phenomena in BaCeO<sub>3</sub>. *Solid State Ion.* **86–88**, 647–652 (1996).
- Münch, W., Kreuer, K. D., Seifert, G. & Maier, J. A quantum molecular dynamics study of proton diffusion in SrTiO<sub>3</sub> and CaTiO<sub>3</sub>. *Solid State Ion.* **125**, 39–45 (1999).
- Shi, J., Zhou, Y. & Ramanathan, S. Colossal resistance switching and band gap modulation in a perovskite nickelate by electron doping. *Nat. Commun.* **5**, 4860 (2014).
- Tuller, H. L. Defect engineering: design tools for solid state electrochemical devices. *Electrochim. Acta* **48**, 2879–2887 (2003).
- Duan, C. *et al.* Readily processed protonic ceramic fuel cells with high performance at low temperatures. *Science* **349**, 1321–1326 (2015).
- Liu, M. & Hu, H. Effect of interfacial resistance on determination of transport properties of mixed-conducting electrolytes. *J. Electrochem. Soc.* **143**, L109–L112 (1996).
- Wang, S., Wu, L., Gao, J., He, Q. & Liu, M. Oxygen ion transference number of doped lanthanum gallate. *J. Power Sources* **185**, 917–921 (2008).
- Steele, B. C. & Heinzel, A. Materials for fuel-cell technologies. *Nature* **414**, 345–352 (2001).
- Ishihara, T., Shibayama, T., Honda, M., Nishiguchi, H. & Takita, Y. Intermediate temperature solid oxide fuel cells using LaGaO<sub>3</sub> electrolyte II. Improvement of oxide ion conductivity and power density by doping Fe for Ga site of LaGaO<sub>3</sub>. *J. Electrochem. Soc.* **147**, 1332–1337 (2000).
- Esposito, V. & Traversa, E. Design of electroceramics for solid oxides fuel cell applications: playing with ceria. *J. Am. Ceram. Soc.* **91**, 1037–1051 (2008).
- Pergolesi, D. *et al.* High proton conduction in grain-boundary-free yttrium-doped barium zirconate films grown by pulsed laser deposition. *Nat. Mater.* **9**, 846–852 (2010).
- Fabbri, E., D'Epifanio, A., Di Bartolomeo, E., Licoccia, S. & Traversa, E. Tailoring the chemical stability of Ba(Ce<sub>0.8</sub>–<sub>x</sub>Zr<sub>x</sub>)Y<sub>0.2</sub>O<sub>3–δ</sub> protonic conductors for intermediate temperature solid oxide fuel cells (IT-SOFCs). *Solid State Ion.* **179**, 558–564 (2008).
- Haile, S. M., Boysen, D. A., Chisholm, C. R. I. & Merle, R. B. Solid acids as fuel cell electrolytes. *Nature* **410**, 910–913 (2001).
- Shim, J. H., Gür, T. M. & Prinz, F. B. Proton conduction in thin film yttrium-doped barium zirconate. *Appl. Phys. Lett.* **92**, 253115 (2008).
- Medarde, M. L. Structural, magnetic and electronic properties of RNiO<sub>3</sub> perovskites (R=rare earth). *J. Phys. Condens. Matter* **9**, 1679 (1997).
- Kreuer, K. On the complexity of proton conduction phenomena. *Solid State Ion.* **136/137**, 149–160 (2000).
- Jaramillo, R., Ha, S. D., Silevitch, D. M. & Ramanathan, S. Origins of bad-metal conductivity and the insulator-metal transition in the rare-earth nickelates. *Nat. Phys.* **10**, 304–307 (2014).
- García, J., Blasco, J., Proietti, M. G. & Benfatto, M. Analysis of the x-ray-absorption near-edge-structure spectra of La<sub>1–x</sub>Nd<sub>x</sub>NiO<sub>3</sub> and LaNi<sub>1–x</sub>Fe<sub>x</sub>O<sub>3</sub> perovskites at the nickel K edge. *Phys. Rev. B* **52**, 15823–15828 (1995).
- Medarde, M. *et al.* Charge disproportionation in RNiO<sub>3</sub> perovskites (R=rare earth) from high-resolution x-ray absorption spectroscopy. *Phys. Rev. B* **80**, 245105 (2009).
- Tan, Z., Heald, S. M., Cheong, S. W., Cooper, A. S. & Moodenbaugh, A. R. Nature of hole doping in Nd<sub>2</sub>NiO<sub>4</sub> and La<sub>2</sub>NiO<sub>4</sub>: Comparison with La<sub>2</sub>CuO<sub>4</sub>. *Phys. Rev. B* **47**, 12365–12368 (1993).
- O'Grady, W. E., Pandya, K. I., Swider, K. E. & Corrigan, D. A. In situ x-ray absorption near-edge structure evidence for quadrivalent nickel in nickel battery electrodes. *J. Electrochem. Soc.* **143**, 1613–1617 (1996).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** Financial support was provided by the Army Research Office (grants W911NF-14-1-0348 and W911NF-14-1-0669), the Air Force Office of Scientific Research (grant FA9550-12-1-0189), the Advanced Research Projects Agency-Energy (ARPA-E), an IBM PhD Fellowship and the National Academy of Sciences. Part of the work was performed at the Center for Nanoscale Systems at Harvard University. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract number DE-AC02-06CH11357. D.D.F. was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division.

**Author Contributions** Y.Z. and S.R. conceived the study. Y.Z. fabricated the fuel cells and performed the initial tests. X.G. designed and performed the quantitative fuel-cell tests and analysis. Y.Z., H.Z., H.L. and S.L. performed the X-ray absorption spectroscopy measurements. Y.Z. and H.Z. conducted the X-ray diffraction characterizations. K.R. performed the low-temperature electronic transport measurements. S.A. prepared the freestanding Si<sub>3</sub>N<sub>4</sub> membrane. M.T. provided technical advice on the micro-SOFC fabrication and characterization. Y.Z., X.G., J.S. and S.R. wrote the manuscript. All authors discussed the results and commented on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.R. ([shriram@purdue.edu](mailto:shriram@purdue.edu)) or Y.Z. ([youzhou@fas.harvard.edu](mailto:youzhou@fas.harvard.edu)).



## METHODS

**Micro-fabricated SOFC.** 4-inch-diameter, 525- $\mu\text{m}$ -thick Si (100) wafers coated with 200 nm  $\text{Si}_3\text{N}_4$  on both sides were used as substrates for micro-fabrication of an SOFC. One side of the wafer was patterned with photolithography to define silicon nitride areas uncovered by photoresist. Then the uncovered silicon nitride was removed by reactive ion etching in  $\text{CF}_4$  and  $\text{O}_2$ . Afterwards, the exposed Si was etched with a 30 wt% KOH aqueous solution at 86 °C for  $\sim 5$  h to leave a  $160 \times 160 \mu\text{m}^2$  free-standing  $\text{Si}_3\text{N}_4$  membrane. After KOH etching, the 4-inch wafers were cut into  $1 \times 1 \text{ cm}^2$  chips, with nine windows on each of the chips. SNO electrolyte films with thicknesses ranging from 50 nm to 1.5  $\mu\text{m}$  were deposited onto the silicon nitride membranes by radio-frequency magnetron sputtering in an  $\text{Ar}/\text{O}_2$  mixture at a total pressure of 5 mTorr, either from a ceramic SNO target, or from two metallic Ni and Sm targets. For films sputtered from metal targets, the chips were annealed under 100 bar of pure  $\text{O}_2$  at 500 °C for 24 h so that SNO would form the perovskite phase after annealing. The growth rate of SNO was calibrated by X-ray reflectivity, cross-section transmission electron microscopy and scanning electron microscopy. The Sm:Ni cation ratio was determined by energy-dispersive X-ray spectroscopy. Then, the Pt cathode was deposited by magnetron sputtering in pure Ar at a total pressure of 75 mTorr, which yields a porous Pt layer to increase the size of the triple phase boundaries on the anode and cathode side. The  $\text{Si}_3\text{N}_4$  layer on the backside was removed by reactive ion etching in  $\text{CF}_4$  and  $\text{O}_2$ . Finally, a Pt anode layer was deposited into the Si well side using sputtering under the same conditions as for the cathode. Dense palladium films of thickness 100–200 nm were also used as the anode, which were deposited by an electron-beam evaporator. The detailed fabrication process is shown in Extended Data Fig. 3.

**Synthesis of SNO epitaxial thin films.** The epitaxial SNO thin films were grown on LAO (001) by radio-frequency magnetron sputtering in an  $\text{Ar}/\text{O}_2$  mixture at total pressure of 5 mTorr from two metallic Ni and Sm targets. The samples were sealed in a vessel under 100 bar of pure  $\text{O}_2$  and annealed at 500 °C for 24 h in a tube furnace.

**Electrical and electrochemical characterization.** Fuel-cell tests were performed in a custom-design fuel cell test station. The morphology of the membranes during fuel cell testing was monitored *in situ* under an optical microscope. Anode current was collected with a gold O-ring and a stainless-steel base, and the cathode current was collected through a micromanipulator probe with a Pt-plated tungsten tip. The electrochemical active area for fuel cell performance was defined as the area of the free-standing SNO membrane. For epitaxial thin films, the conductivity measurements were done using in-plane geometry with porous Pt electrodes. The current–voltage characteristics were measured by starting at an OCV and sweeping down to 0 V at a rate of 20  $\text{mV s}^{-1}$  (or 10  $\text{mV s}^{-1}$ ). Electrochemical impedance spectroscopy was scanned from  $10^6$  Hz to 1 Hz with an amplitude of 20 mV. All the electrochemical measurements were performed with a Solartron 1260/1287 electrochemical test setup. The impedance data were fitted using ZView software. For Pt/SNO/Pt fuel cells, either dry or moist 5%  $\text{H}_2/95\%$  Ar was flown onto the anode side. For Pt/SNO/Pd fuel cells, pure  $\text{H}_2$  bubbled through room-temperature water was flown onto the Pd anode. In both cases, stationary air was used as the cathode oxidant. The ionic conductivity of epitaxial thin films was measured in

dry 5%  $\text{H}_2/95\%$  Ar and the conductivity of suspended membranes in Pt/H-SNO/Pt SOFCs was measured with 3% humidified 5%  $\text{H}_2/95\%$  Ar as fuel and laboratory air as oxidant.

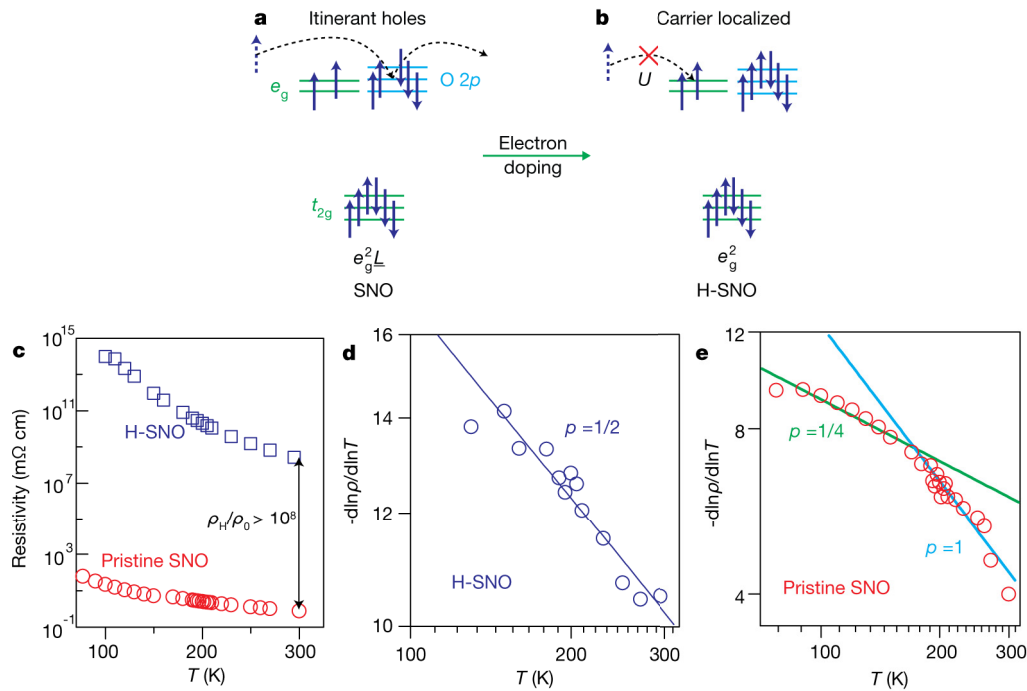
*In situ* conductivity dynamics measurements were performed with a Keithley 2635A and Solartron 1260/1287 in a custom-built chamber by switching between dry 5%  $\text{H}_2/95\%$  Ar and  $\text{O}_2$  with a fixed flow rate of 150 standard cubic centimetres per minute (sccm). Electronic transport studies below room temperature were performed in vacuum *ex situ* using a Lakeshore probe-station and a Keithley 2635A on samples annealed in dry 5%  $\text{H}_2/95\%$  Ar for 30 min at 200 °C.

**X-ray absorption spectroscopy studies.** The X-ray absorption spectroscopy data were acquired at the bending magnet beamline, 12-BM-B, at the Advanced Photon Source, Argonne National Laboratory. The absorption was measured in fluorescence mode with the samples placed in a custom-made cell allowing *in situ* control of the atmosphere and heating of the sample. An infrared heater is used to heat the sample up to 200 °C. A 13-element Ge detector (Canberra) was used to measure the fluorescence yield. Grazing incidence geometry was used to minimize the elastic scattering intensity. The incident angle is varied from 0.25° to 5°, covering a range below and above the critical angle. The calibration of the monochromator was monitored by simultaneously measuring the absorption of a nickel reference foil during each measurement. For *ex situ* XANES measurements, SNO samples annealed in dry 4%  $\text{H}_2/96\%$  Ar for 30 min at 200 °C or 300 °C. The data were normalized by fitting the pre-edge to zero and the post-edge to 1 using Ifeffit performed by the software Athena (<http://cars9.uchicago.edu/ifeffit/Ifeffit>). Both epitaxial SNO thin films of different thickness on LAO and polycrystalline SNO thin films on  $\text{SiO}_2/\text{Si}$  were characterized by XANES.

**Synchrotron X-ray diffraction.** Synchrotron X-ray diffraction of the SNO samples were conducted at an insertion device beamline, 12ID-D at the Advanced Photon Source on a six-circle Huber goniometer with an X-ray energy of 20 keV using a pixel array area detector (Dectris Pilatus 100 K). The X-ray beam had a flux of  $10^{12}$  photons per second. The  $q_z$ -scan (L-scan) was obtained by removing the background scattering contributions using the two-dimensional images. For *ex situ* X-ray diffraction measurements, SNO samples were grown on LAO substrates and annealed in 5%  $\text{H}_2/95\%$  Ar at 300 °C for 2 h. For the real-space mapping shown in Extended Data Fig. 9d–f, an X-ray footprint of 50  $\mu\text{m}$  (horizontal in Extended Data Fig. 9b)  $\times$  500  $\mu\text{m}$  (vertical in Extended Data Fig. 9b) was used to scan across the sample, collecting the diffraction pattern from each point.

**SNO stability test.** To test the material stability in a pure hydrogen atmosphere, we annealed SNO thin films under 1 bar of pure  $\text{H}_2$  in a tube furnace at 500 °C for 48 h and 72 h. Pt electrodes were deposited onto SNO thin films as catalyst. The  $\text{H}_2$  flow was set to a constant of 300 sccm. X-ray diffraction was performed on the annealed samples using a Bruker-D8 Discover diffractometer.

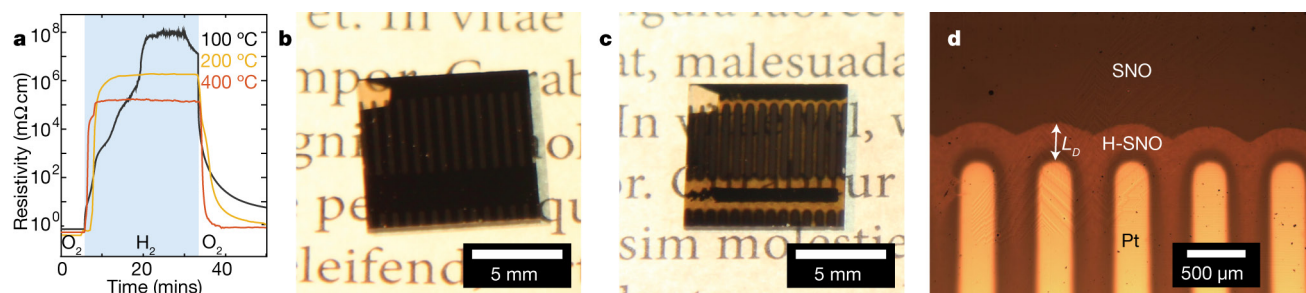
- Shklovskii, B. I. & Efros, A. L. *Electronic Properties of Doped Semiconductors* Ch. 9/10, 202–250 (Springer, 1984).
- Goodenough, J. B. Electronic and ionic transport properties and other physical aspects of perovskites. *Rep. Prog. Phys.* **67**, 1915 (2004).
- Natoli, C. R. in *EXAFS and Near Edge Structure III* Vol. 2 *Springer Proceedings in Physics* (eds Hodgson, K. O., Hedman, B. & Penner-Hahn, J. E.) Ch. 10, 38–42 (Springer, 1984).



**Extended Data Figure 1 | The electronic structure of SNO and H-SNO in the covalent limit and their electronic transport mechanisms.**

**a**, The electronic structure of SNO in the covalent limit. Ligand holes are present on O  $2p$  orbitals of the pristine SNO, while two electrons occupy the Ni  $e_g$  manifold. The pristine SNO, however, is not strongly correlated because carriers transport through the O  $2p$  ligand holes. **b**, The electronic structure of H-SNO. Upon electron doping and thus filling of the ligand holes, electrons have to overcome Hubbard intra-orbital correlation  $U$  to transport, which opens up a large Mott gap, and suppresses the electronic conduction in SNO. **c**, The resistivity  $\rho$  of H-SNO compared with pristine

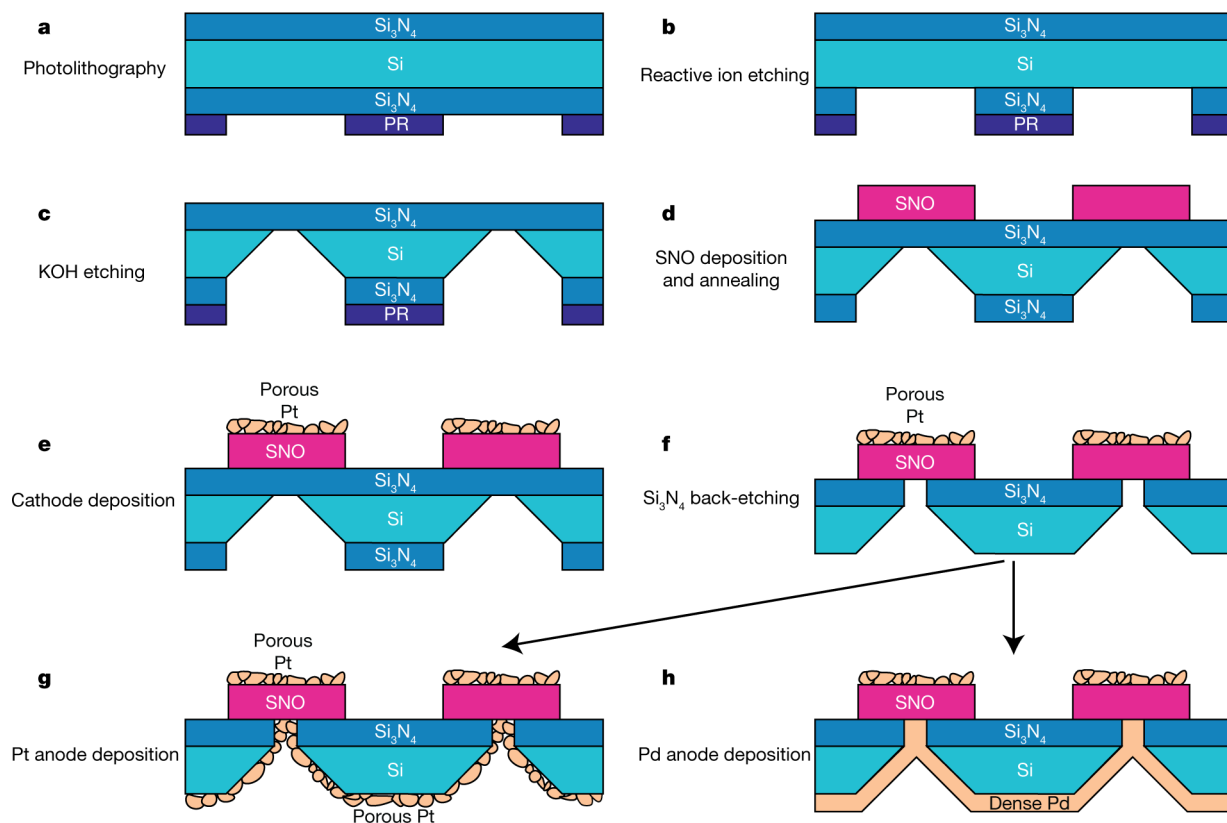
SNO. The resistivity of H-SNO is more than eight orders of magnitude larger than that of pristine SNO at room temperature. **d**, **e**, Derivatives of resistivity ( $-\mathrm{d}\ln\rho/\mathrm{d}\ln T$ ) as a function of  $T$  plotted in log-log scale for H-SNO and SNO. The transport mechanism can be determined from the slope  $p$  of the  $-\mathrm{d}\ln\rho/\mathrm{d}\ln T$  versus  $T$  curves. H-SNO shows the Efros-Shklovskii variable range hopping mechanism ( $p = 1/2$ ), indicating polaron formation in the presence of a Coulomb gap<sup>31</sup> (**d**). Pristine SNO shows crossover from activated conduction ( $p = 1$ ) to Mott variable range hopping ( $p = 1/4$ ) (**e**). The Coulomb repulsion is less strong in pristine SNO.



**Extended Data Figure 2 | Fuel-induced suppression of electronic conduction in SNO.** **a**, Temporal evolution of SNO conductivity when switching between different gas environments at various temperatures. **b–d**, Images of SNO and H-SNO on transparent substrate LAO. **b**, Pristine SNO shows dark, shining colour and the Pt bars are bright. **c**, After annealing in 5% H<sub>2</sub>/95% Ar at 300 °C for 1.5 h and cooling down to room

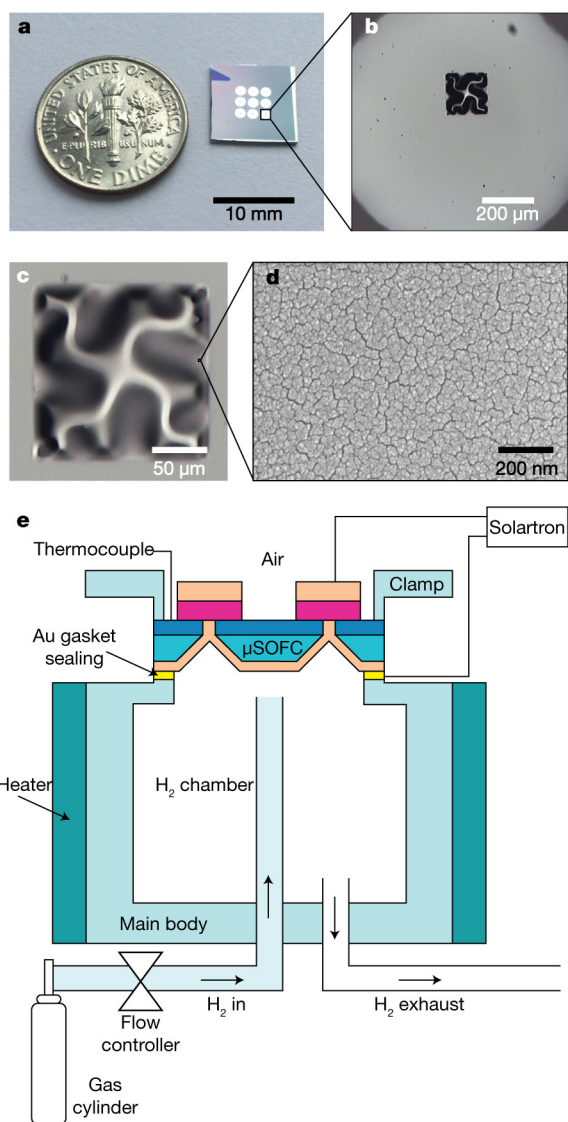
temperature in the same gas environment, SNO near the Pt electrodes becomes electronically insulating and transparent. A clear diffusion profile can be seen as the transparent region has a shape similar to the outline of the Pt electrodes. **d**, An optical micrograph of the hydrogenated SNO indicates a diffusion profile of protons from the triple phase boundaries. The diffusion length  $L_D$  is estimated to be  $\sim 300 \mu\text{m}$ .



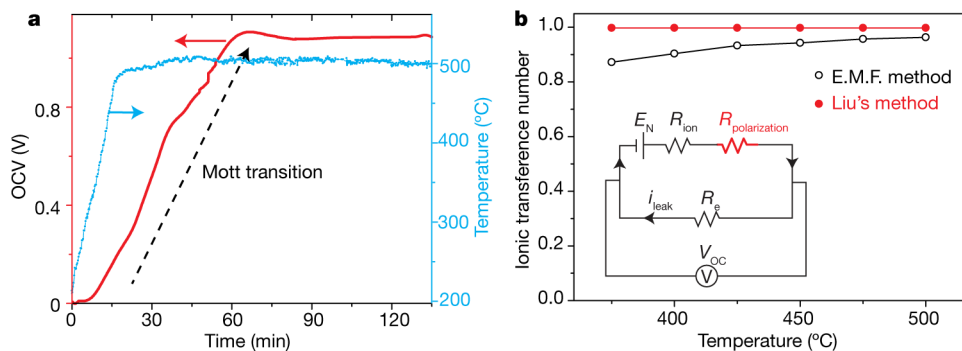


**Extended Data Figure 3 | A schematic of the fabrication process of fuel cells with free-standing SNO membranes as the electrolyte.** **a**, Patterning a etch mask on the back side of the  $\text{Si}_3\text{N}_4/\text{Si}/\text{Si}_3\text{N}_4$  chip with photoresist (PR). **b**, Removing exposed silicon nitride by reactive ion etching in  $\text{CF}_4$  and  $\text{O}_2$ . **c**, Etching the Si from the back side with a KOH aqueous solution to make free-standing  $\text{Si}_3\text{N}_4$  membrane. **d**, Depositing SNO thin films onto the  $\text{Si}_3\text{N}_4$  membranes by radio-frequency magnetron sputtering and post-annealing the sample to form stoichiometric SNO. **e**, Fabricating the

porous Pt cathodes on the front side of the chip. **f**, Removing the silicon nitride membrane from the back side of the chip to expose SNO, using reactive ion etching. **g**, **h**, Depositing anodes on the back side of the chip. Two types of fuel cell anodes were studied in this work: porous Pt as a model system (**g**) and a dense Pd anode (**h**). Pd is an industry-standard proton conducting membrane that is used in this study to selectively permeate protons from the fuel side to the cathode.



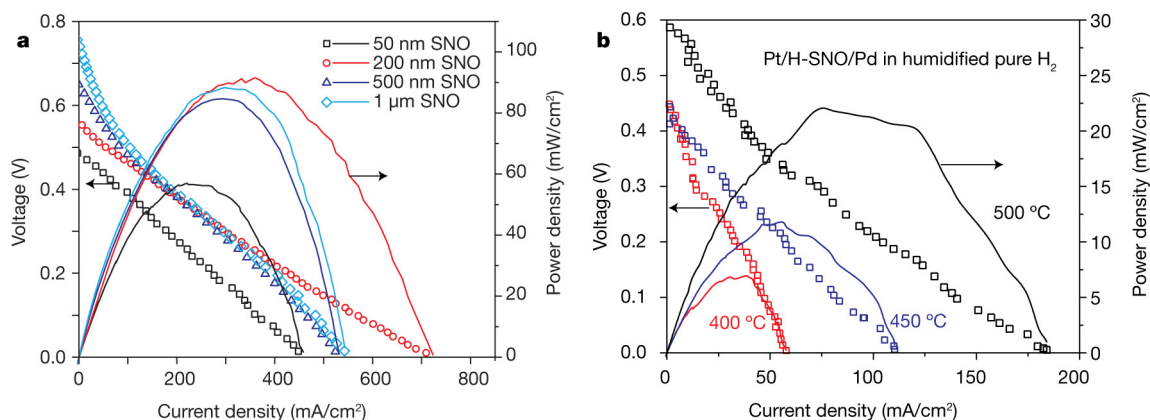
**Extended Data Figure 4 | SNO micro-fabricated SOFCs and fuel cell test apparatus.** **a**, An image of a 10 mm × 10 mm  $\text{Si}_3\text{N}_4/\text{Si}$  chip with nine SNO-electrolyte fuel cells (US dime coin shown for size). **b**, **c**, Optical micrographs of the free-standing buckled SNO membrane due to local compressive strain with top Pt cathode on a Si chip. The buckled morphology is due to local compressive strain, engineered intentionally by synthesis and is critical for the mechanical stability and performance of the SOFC. **d**, A scanning electron microscope of the top porous Pt electrode. **e**, A schematic of the customized low-temperature micro-fabricated SOFC ( $\mu\text{SOFC}$ ) testing station. Both pure  $\text{H}_2$  and 5% $\text{H}_2$ /95% Ar were used as fuel in the experiments.



**Extended Data Figure 5 | OCV of H-SNO micro-fabricated SOFCs.** **a**, Temporal evolution of the OCV of a Pt/H-SNO/Pt micro-fabricated SOFC with a 3% humidified 5%  $H_2$ /95% Ar as fuel and laboratory air as oxidant, as the temperature ramps up. Initially SNO is electronically conductive so the OCV is close to zero. During the hydrogenation process, the OCV continues to increase after the temperature is stabilized and reaches near-ideal OCV, indicating that electronic conduction is almost completely suppressed in H-SNO by the Mott transition. The hydrogen fuel was always supplied at a constant flow rate both before  $t = 0$  and during the experiments, and the initial low OCV is not due to the lack of fuel. **b**, The ionic transference number of H-SNO at various temperatures of interest to low-temperature SOFCs measured in Pt/SNO/Pt cells. Two methods can be used to calculate the ionic transference number. In the electromotive force (E.M.F.) method, the fuel cell under the OCV condition (infinitely large external resistance load) is modelled with an equivalent circuit containing a voltage source with an output voltage of

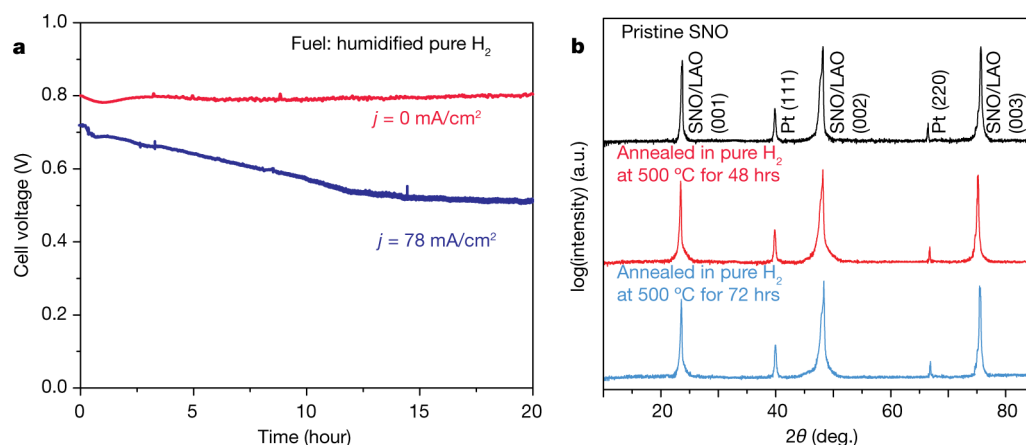
Nernst potential  $E_N$ , and two resistors  $R_{ion}$  and  $R_e$ , which correspond to the electrolyte's ionic resistance and electronic resistance, respectively. The equivalent circuit is similar to the one shown in the inset, but without the  $R_{polarization}$  element (drawn in red).  $V_{OC}$  is the measured OCV. Note that there will be a small leakage current  $i_{leak}$  due to the finite electronic resistance of the electrolyte, but the electromotive force method assumes that the interface processes are infinitely fast and omits the polarization loss. In the method developed by Liu *et al.*<sup>15</sup>, since there is a very small leakage electronic current flowing through the electrolyte, one needs to consider the electrode polarization loss. Therefore, an extra resistive element ( $R_{polarization}$ ) needs to be considered in the equivalent circuit as shown in the inset (the red-coloured element corresponds to the extra term). With reduced polarization and increased electrolyte resistance, the ionic transference number calculated by the two methods tends to converge (see Supplementary Information for more discussion).





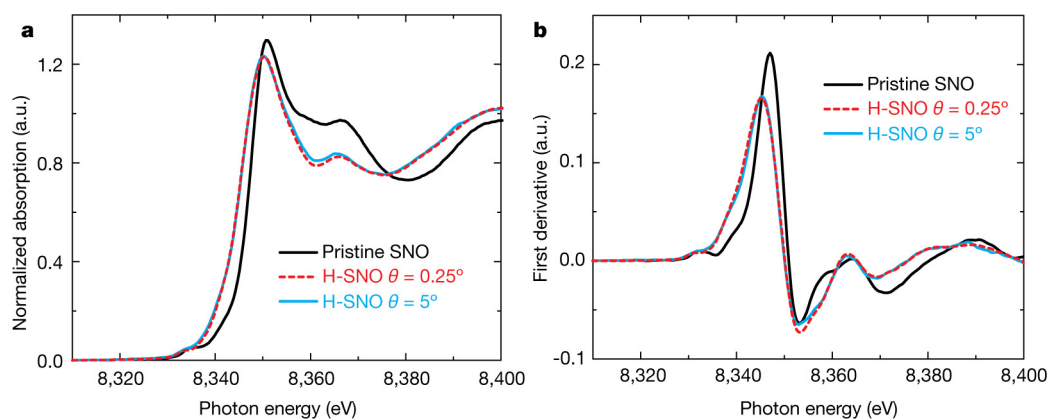
**Extended Data Figure 6 | H-SNO fuel cell performance. a,** The dependence of micro-fabricated SOFC performance on the thickness of the SNO electrolyte at 500 °C. We fabricated a series of samples with various thicknesses of the electrolyte while keeping identical deposition conditions for the cathode and anode. By doing so, the electrolyte Ohmic resistance is varied while the electrode polarization resistance is kept more or less a constant. A clear increase in OCV with increasing thickness can be seen, which could be due to the decrease in the electrode polarization loss because of the larger electrolyte Ohmic resistance, as discussed in Extended Data Fig. 5. The power density does not show much dependence on the electrolyte thickness, because thicker electrolytes leads to higher Ohmic resistance, but also higher OCV. **b,** Performance of Pt/SNO/Pd micro-fabricated SOFCs with a dense Pd anode with 3% humidified pure H<sub>2</sub> as fuel and laboratory air as oxidant. It has been shown that hydrogen

primarily creates protonic defects rather than oxygen vacancies in SNO (ref. 12). To verify that protons are the dominant mobile ion species in SNO and H-SNO, we fabricated an SOFC with the SNO electrolyte, a dense 100-nm-thick Pd anode, and a porous 100-nm-thick Pt cathode. Pd anode is known as a protonic conductor but an oxygen ion barrier and can therefore filter out any oxygen ion transport. This verifies that protons rather than oxygen ions are the dominant mobile ions in SNO. During the fuel cell testing, 100 sccm pure H<sub>2</sub> was flowed on the anode side, with the cathode exposed to air. The fuel cell with dense Pd has an OCV of 0.6 V and a peak power density of 24 mW cm<sup>-2</sup> at 500 °C. The protonic conductivity of H-SNO can be extrapolated from impedance spectroscopy and OCV measurements. The similar values of the measured ionic conductivity in cells with Pt and Pd anode confirm that protonic conduction is the dominant ionic transport mechanism.



**Extended Data Figure 7 | Stability of H-SNO.** **a**, Cell voltage measured at 500 °C for a Pt/SNO/Pd fuel cell with wet 100%  $H_2$  as the fuel and stationary air as oxidant with current being  $0 \text{ mA cm}^{-2}$  (OCV condition) and  $78 \text{ mA cm}^{-2}$ , respectively. The operation is stable for more than 20 h, implying that H-SNO exhibits considerable stability for fuel cell operation. The power output decreases slightly as a function of time

owing to coarsening-induced porosity reduction of the metallic electrodes when current is drawn at 500 °C. **b**, X-ray diffraction pattern of SNO, and H-SNO (on LAO substrates) after being annealed under 1 bar of pure  $H_2$  at 500 °C for 48 h and 72 h. No new diffraction peaks are observed after annealing, which shows that H-SNO is quite stable in pure  $H_2$  for extended periods of time.  $\theta$  is the incident angle of the X-ray.

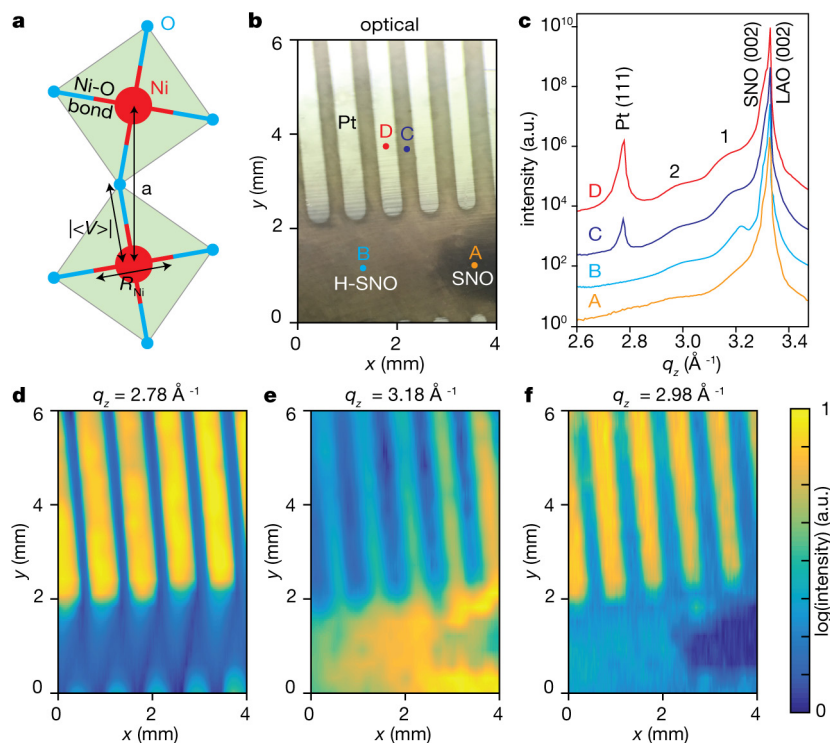


### Extended Data Figure 8 | Angle-dependent XANES characterization.

**a**, *Ex situ* angle-dependent XANES spectra of hydrogenated SNO with a reference spectrum from pristine SNO. The critical angle  $\theta_c$  of X-ray scattering for SNO at the X-ray energy near Ni K-edge is calculated to be  $0.335^\circ$ . When the X-ray incident angle is below the critical angle ( $0.25^\circ$ ), the XANES signal is surface sensitive with a penetration depth of  $\sim 10$  nm. For an incident angle of  $5^\circ$ , the penetration depth is close to

$1\ \mu\text{m}$ . The absence of angle-dependence of the XANES spectra shows that the hydrogen incorporation happens almost homogeneously across the film thickness. The XANES spectrum acquired at incident angle of  $1^\circ$  (not shown) is also similar to those at  $0.25^\circ$  and  $5^\circ$ . **b**, The first derivative of the normalized absorption shows a similar change in the average valence state of Ni at the film surface and in the bulk.

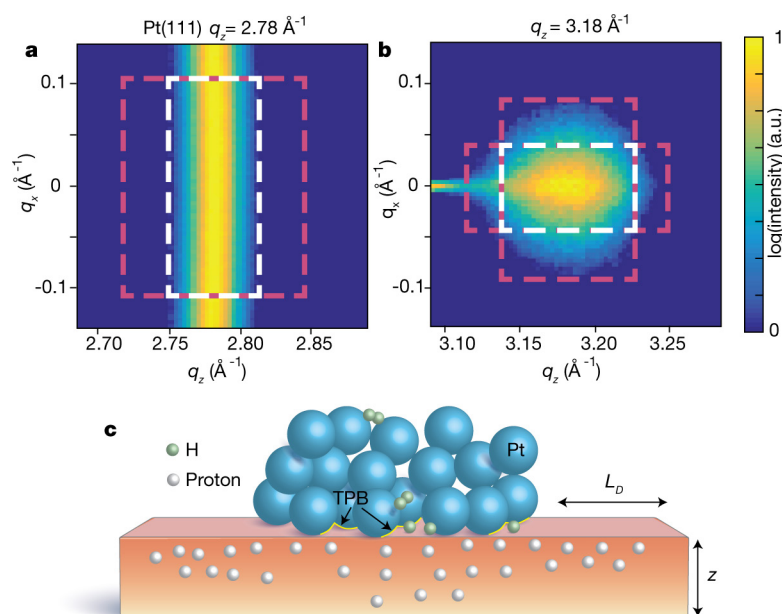




### Extended Data Figure 9 | Synchrotron structural characterization of the emergent SNO phase.

**a**, An increase in the lattice constant can be caused by the larger crystal radius of  $\text{Ni}^{2+}$  and electron localization. When the formal valence state of Ni reduces, its ionic radius  $R_{\text{Ni}}$  increases, leading to the elongation of the Ni–O bond. In addition to the simple valence-state-related lattice expansion, electron localization can also increase the metal–oxygen bond length, which can be understood on the basis of the virial theorem for central-force fields:  $2\langle T \rangle + \langle V \rangle = 0$ , where  $\langle T \rangle$  is the mean kinetic energy of electrons, and  $\langle V \rangle$  is the average potential energy. When transitioning from itinerant to localized electronic behaviour, the absolute value  $|\langle V \rangle|$  must decrease, which is achieved by a longer metal–oxygen bond length<sup>32</sup>, that is,  $\langle \text{Ni–O} \rangle_{\text{loc}}$  exceeds  $\langle \text{Ni–O} \rangle_{\text{itin}}$  even for the same valence state. **b**, An optical image of a hydrogenated SNO sample. H-SNO phase forms near and under the Pt electrodes, while a part of the sample remains in its pristine phase. **c**, X-ray diffraction patterns from the various spots A, B, C and D marked in **b**. The SNO and LAO peaks are indexed in pseudocubic notation. As the pristine SNO has a pseudocubic lattice constant close to that of the LAO, the SNO (002) appears almost as a shoulder of the LAO (002) peak. With decreasing distance between the

X-ray spot and Pt electrodes, SNO (002) indeed shifts to smaller  $q_z$  (no other peaks observed). Two peaks (peak 1 at  $q_z = 3.18 \text{ \AA}^{-1}$  and peak 2 at  $q_z = 2.98 \text{ \AA}^{-1}$ ) appear in the hydrogenated region and correspond to  $\sim 4\%$  and  $\sim 10\%$  increase in the lattice constant. Peak 2 has the largest intensity right underneath the Pt catalyst, while peak 1 has the highest intensity far away from the Pt electrodes. The difference in the lattice constant change can be related to the decreasing doping concentration with increasing diffusion length from the triple phase boundary where hydrogen enters SNO (Extended Data Fig. 10c). **d–f**, Real-space mapping of the intensity of the Pt (111) peak at  $q_z = 2.78 \text{ \AA}^{-1}$  (**d**), the H-SNO peak 1 at  $q_z = 3.18 \text{ \AA}^{-1}$  (**e**) and the H-SNO peak 2 at  $q_z = 2.98 \text{ \AA}^{-1}$  (**f**). A clear positive correlation between the Pt (111) and the  $q_z = 2.98 \text{ \AA}^{-1}$  peaks can be seen, whereas the Pt (111) and  $q_z = 3.18 \text{ \AA}^{-1}$  peaks show a negative correlation. The intensity of both peaks 1 and 2 is low in the pristine region, as expected. The increase in the average Ni–O bond length can be also inferred from XANES spectra using Natoli's rule<sup>33</sup>, which states that the energy separation between features B, D, and E will scale inversely with the square of the Ni–O distance, because they are derived from the first oxygen coordination cell<sup>27</sup>.



**Extended Data Figure 10 | Raw X-ray diffraction patterns and a schematic of proton diffusion.** **a, b,** The collected raw two-dimensional diffraction patterns for the real-space mapping in Extended Data Fig. 9. To get the real-space mapping of different peaks across the sample, we scan the sample with an X-ray footprint of  $50\ \mu\text{m}$  (horizontal in Extended Data Fig. 9b)  $\times$   $500\ \mu\text{m}$  (vertical in Extended Data Fig. 9b) by collecting the diffraction pattern from each point with an area detector. Then we calculate the diffraction intensity of each peak (Pt (111), and peak 1, 2 in Extended Data Fig. 9c) at each real-space spot from the 2d images and map it into real space to create Extended Data Fig. 9d–f. **a,** Diffraction pattern of Pt (111) from a spot on the Pt electrode. A diffraction ring is observed as Pt is polycrystalline. **b,** Diffraction pattern at  $q_z = 3.18\ \text{\AA}^{-1}$  from a spot between the Pt electrodes. Unlike the Pt pattern, it shows up as a point with a truncation rod rather than a ring in  $k$ -space, indicating that H-SNO is still epitaxial on LAO after hydrogenation. For both **a** and **b** the

region inside the white dashed line was used to calculate the signal, while the region enclosed by the red dashed line but not by the white dashed line was used to calculate the background along both the  $q_z$  and  $q_x$  directions. The signal/background region and calculation algorithm were kept the same for all the real-space spots measured on the sample for a particular spot in the reciprocal space. **c,** A schematic of proton incorporation and diffusion near Pt electrodes. The part of SNO directly underneath the porous Pt electrodes is on average closer to the triple phase boundaries (TPBs) than the SNO region between the Pt electrodes. Therefore, a higher concentration of protons is expected under the Pt electrodes, which explains the larger lattice constant change and the correlation relation found in Extended Data Fig. 9. As the thickness of the film ( $z \sim 100\ \text{nm}$ ) is much smaller than the diffusion length (hundreds of micrometres), the proton concentration should not vary much along the thickness direction for the case of epitaxial thin films on LAO.

# An autonomous chemically fuelled small-molecule motor

Miriam R. Wilson<sup>1</sup>, Jordi Solà<sup>1\*†</sup>, Armando Carlone<sup>1\*†</sup>, Stephen M. Goldup<sup>1†</sup>, Nathalie Lebrasseur<sup>1†</sup> & David A. Leigh<sup>1</sup>

Molecular machines are among the most complex of all functional molecules and lie at the heart of nearly every biological process<sup>1</sup>. A number of synthetic small-molecule machines have been developed<sup>2</sup>, including molecular muscles<sup>3,4</sup>, synthesizers<sup>5,6</sup>, pumps<sup>7–9</sup>, walkers<sup>10</sup>, transporters<sup>11</sup> and light-driven<sup>12–16</sup> and electrically<sup>17,18</sup> driven rotary motors. However, although biological molecular motors are powered by chemical gradients or the hydrolysis of adenosine triphosphate (ATP)<sup>1</sup>, so far there are no synthetic small-molecule motors that can operate autonomously using chemical energy (that is, the components move with net directionality as long as a chemical fuel is present)<sup>19</sup>. Here we describe a system in which a small molecular ring (macrocycle) is continuously transported directionally around a cyclic molecular track when powered by irreversible reactions of a chemical fuel, 9-fluorenylmethoxycarbonyl chloride. Key to the design is that the rate of reaction of this fuel with reactive sites on the cyclic track is faster when the macrocycle is far from the reactive site than when it is near to it. We find that a bulky pyridine-based catalyst promotes carbonate-forming reactions that ratchet the displacement of the macrocycle away from the reactive sites on the track. Under reaction conditions where both attachment and cleavage of the 9-fluorenylmethoxycarbonyl groups occur through different processes, and the cleavage reaction occurs at a rate independent of macrocycle location, net directional rotation of the molecular motor continues for as long as unreacted fuel remains. We anticipate that autonomous chemically fuelled molecular motors will find application as engines in molecular nanotechnology<sup>2,19,20</sup>.

The design of nanometre-scale motors in which the components incessantly rotate with net directionality has tantalized scientists since Feynman's discussion of the physics of a theoretical tiny ratchet-and-pawl<sup>21</sup>. In the 1990s Kelly's group produced a series of molecular analogues of a ratchet-and-pawl, confirming the lack of directional bias in the movement of the components at equilibrium<sup>22</sup>. Their designs culminated in a system that employed chemical reactions to bias a 120° rotation of a triptycene residue in one direction<sup>23</sup>, but attempts to extend this approach to repetitive 360° directional rotation proved unsuccessful<sup>24</sup>. Light-driven rotary molecular motors based on overcrowded alkenes<sup>12,13</sup> and imines<sup>14,16</sup> have been developed by the groups of Feringa and Lehn, while our group<sup>25,26</sup> and others<sup>27–29</sup> have made molecules in which the components can be rotated directionally stepwise by repetitively carrying out several chemical reactions in sequence. The latter systems all operate through Brownian ratchet mechanisms, differentiating the rates of random thermal motion of the components in each direction by the manipulation of kinetic (mainly steric) barriers<sup>20</sup>. Autonomous operation requires the ratchet mechanism to operate continuously, meaning that the barriers must be repeatedly raised and lowered under the same set of reaction conditions and coupled to the consumption of a chemical species in order to avoid falling foul of the second law of thermodynamics<sup>30</sup>.

The structure and mechanism of operation of a rotary molecular motor (**1**) that continuously rotates its components with net directionality when driven by chemical energy is shown in Fig. 1. The molecule is a [2]catenane featuring two interlocked molecular rings of different sizes. Fumaramide residues (shown in green) on the larger ring (the 'track') serve as binding sites for a smaller benzylic amide macrocycle (blue). Removable bulky groups (red) block the passage of the small ring and, when both blocking groups are attached, trap it in one or other compartment of the cyclic track. As previously demonstrated<sup>31,32</sup>, the macrocycle can be directionally transported between adjacent compartments of a rotaxane thread using the acylation of hydroxyl groups as the energy input. We reasoned that the issue of repeatedly raising and lowering the kinetic barriers to transport under a single set of reaction conditions could be achieved by using a blocking group that attaches and detaches through dissimilar reaction mechanisms: one reaction (for example, attachment) proceeding at rates that vary according to the position of the small macrocycle, the other (for example, cleavage) occurring at a rate independent of the small macrocycle position (an 'information ratchet' mechanism<sup>7,20,31,32</sup>).

In [2]catenane **2/2'**, in which one 9-fluorenylmethoxycarbonyl group of **1** has been cleaved, there is a substantial difference between the distances of the small-ring binding sites (fumaramide groups) and the revealed hydroxyl group; one is very close, where the presence of the ring should inhibit nucleophilic attack by the OH group on a large electrophile, and one too far away for a bound ring to influence rates of reaction noticeably. This should result in dissimilar reaction rates for when the macrocycle occupies the fumaramide unit near ( $k_{\text{close-attach}}$ ) or far from ( $k_{\text{far-attach}}$ ) the hydroxyl group. We carried out model studies on a number of potential chemical fuels, eventually selecting 9-fluorenylmethoxycarbonyl chloride (Fmoc-Cl) because its mechanism of attachment to the molecular motor is very different to that of cleavage of the resulting fluorenylmethoxycarbonate group (shown for the rotaxane model system in Fig. 2). The former occurs by nucleophilic attack of a hydroxyl group directly on the C=O of the chloroformate residue, where the presence or absence of the bulky benzylic amide macrocycle on the adjacent fumaramide group would be expected to influence the reaction rate ( $k_{\text{far-attach}} \neq k_{\text{close-attach}}$ ). In contrast, the detachment reaction occurs by a reaction cascade (eliminating CO<sub>2</sub> and dibenzofulvene) initiated by base abstraction of a proton from the fluorenyl methine group. This is five bonds remote from the site of attachment to the [2]catenane and so the influence of the position of the macrocycle on the detachment reaction rate should be minimal ( $k_{\text{far-cleave}} \approx k_{\text{close-cleave}}$ ). The reactions that lead to the attachment and to the cleavage of the Fmoc group can both be promoted under basic conditions.

Starting from the mono-hydroxyl species **2** and **2'**, Fmoc attachment to **2** should favour formation of the positional isomer *FumD*<sub>2</sub>-**1**

<sup>1</sup>School of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, UK. <sup>†</sup>Present addresses: Institute of Advanced Chemistry of Catalonia (IQAC-CSIC), Jordi Girona 18-26, 08034 Barcelona, Spain (J.S.); Chirotech Technology Centre, Dr Reddy's, Cambridge, CB4 0PE, UK (A.C.); Department of Chemistry, University of Southampton, Southampton SO17 1BJ, UK (S.M.G.); School of Biological and Chemical Sciences, Queen Mary University of London, London E1 4NS, UK (N.L.).

\*These authors contributed equally to this work.

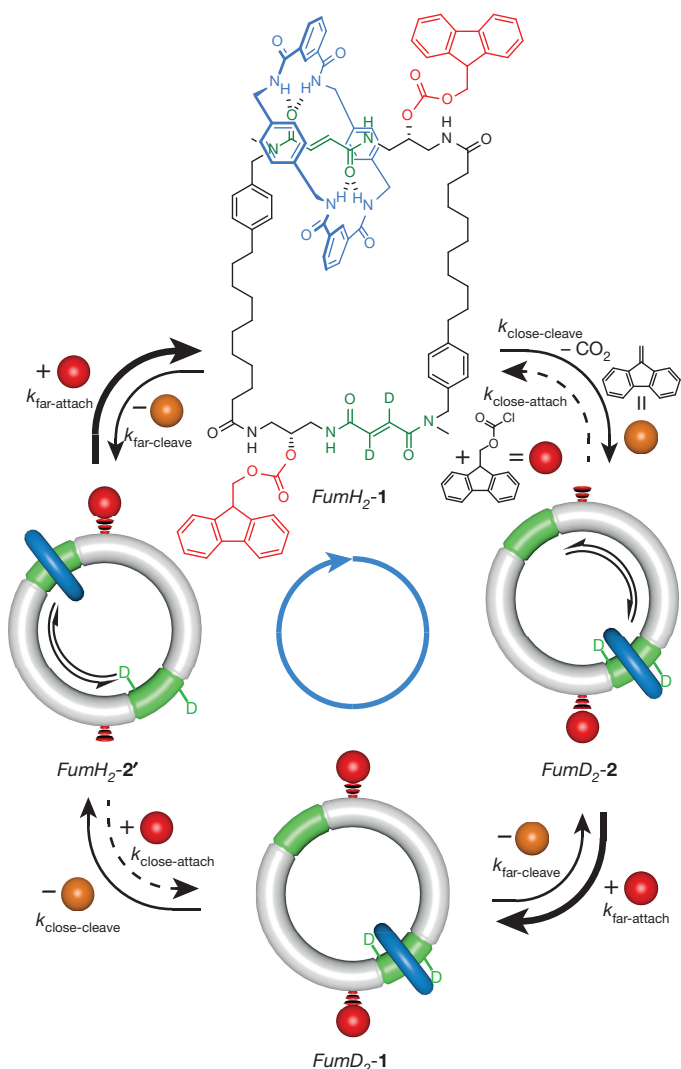


(carbonate formation preferentially occurring distant to the small ring) and likewise Fmoc attachment to 2' should preferentially form *FumH<sub>2</sub>*-1, each reaction causing net transport of the benzylic amide macrocycle in a clockwise direction. The cleavage of either of the Fmoc groups of 1 then occurs (to form 2 or 2' in equal amounts), allowing another Fmoc attachment step to occur, again proceeding with net directional movement of the small ring. To maximize the efficiency of the process, sufficient Fmoc-Cl needs to be present for the attachment reaction to proceed rapidly whenever a hydroxyl group is unmasked. This prevents accumulation of the catenane diol, in which both Fmoc groups have been cleaved, leaving the small ring free to shuttle around the track without directional bias. A fuller discussion of the kinetics<sup>33</sup> of the information ratchet mechanism, including deriving the net directionality of ring rotation from the rate equations, is given in Supplementary Information, section 6.

We first developed the chemistry necessary for the operation of 1 on a simpler [2]rotaxane (a ring threaded on a dumbbell-shaped axle) system, 3 (Fig. 2). [2]Rotaxane 3 was prepared from (*R*)-3-amino-1,2-propanediol (see Supplementary Information, sections 1.2.1 and 1.3.1). When rotaxane 3 was treated with Fmoc-Cl in the presence of a bulky carbonate-forming catalyst, (*R*)-5 (Fig. 2a), the macrocycle was predominantly trapped in the *FumD<sub>2</sub>* compartment (up to 17:83 *FumH<sub>2</sub>*-4:*FumD<sub>2</sub>*-4, as shown by <sup>1</sup>H nuclear magnetic resonance (NMR) spectroscopy in Fig. 2b. Other reaction conditions led to poorer discrimination between the compartments). This result confirms that catalyst (*R*)-5, in its acylated intermediate form (Fig. 2a), can distinguish between the two positional isomers of the rotaxane that interconvert through the macrocycle shuttling between the two fumaramide residues, and that the acylated intermediate preferentially reacts with the hydroxyl group when the macrocycle is on the *FumD<sub>2</sub>* group, that is,  $k_{\text{far-attach}} > k_{\text{close-attach}}$ . Although a chiral catalyst (and chiral motor) was used, mainly for synthetic convenience, the positional bias of the Fmoc addition is almost independent of catalyst handedness (see Supplementary Information, section 5) and stems from one macrocycle binding site being close to the site of reaction on the axle, and the other far away.

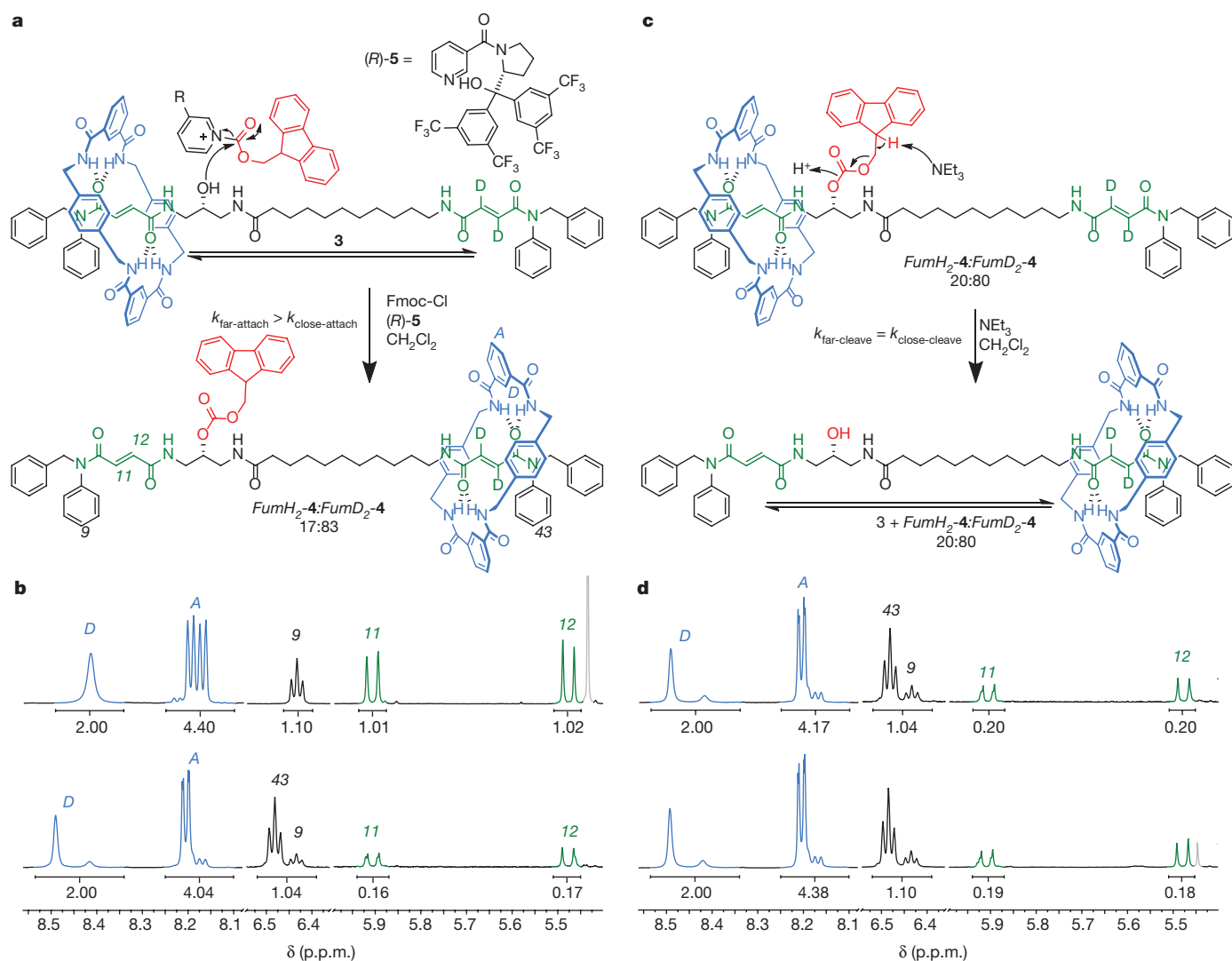
With a directional bias established for the Fmoc addition step, we next investigated the Fmoc cleavage reaction. A solution of 20:80 *FumH<sub>2</sub>*-4:*FumD<sub>2</sub>*-4, in dichloromethane (CH<sub>2</sub>Cl<sub>2</sub>) was treated with triethylamine (NEt<sub>3</sub>) (Fig. 2c). The reaction was sampled at various times, and before all the rotaxane Fmoc groups had been cleaved <sup>1</sup>H NMR analysis of the recovered rotaxane 4 showed the ratio of *FumH<sub>2</sub>*:*FumD<sub>2</sub>* to be unchanged from the starting ratio (for example, rotaxane 4 after 67% formation of 3, Fig. 2d). Thus the Fmoc groups are cleaved from *FumH<sub>2</sub>*-4 and *FumD<sub>2</sub>*-4 at the same rate; the position of the macrocycle in rotaxane 4 does not influence the rate of Fmoc cleavage, that is,  $k_{\text{far-cleave}} = k_{\text{close-cleave}}$ .

Next, conditions were established under which both the Fmoc attachment and cleavage reactions take place in the same reaction mixture (see Supplementary Information, section 2). In a typical procedure, the rotaxane (3 or 4) and (*R*)-5 were dissolved in CH<sub>2</sub>Cl<sub>2</sub> and KHCO<sub>3</sub> was added (to regenerate NEt<sub>3</sub> from hydrochloride salts formed by the cleavage reaction). Solutions of the Fmoc-Cl fuel and Et<sub>3</sub>N in CH<sub>2</sub>Cl<sub>2</sub> were mixed together initially and then more Fmoc-Cl slowly and continuously added using a syringe pump for as long as the motor was required to run. Subjecting rotaxane 4 with an initial macrocycle distribution of 100:0 *FumH<sub>2</sub>*:*FumD<sub>2</sub>* to these operation conditions resulted in 4 with a distribution of 17:83 *FumH<sub>2</sub>*:*FumD<sub>2</sub>* at the steady state (Supplementary Figs 3 and 4). That the Fmoc formation and cleavage reactions run concurrently was further confirmed by showing that a deuterium (D)-labelled Fmoc group on the rotaxane could be exchanged for an unlabelled one under these operating conditions. Treatment of D<sub>2</sub>-(33:67 *FumH<sub>2</sub>*:*FumD<sub>2</sub>*)-4 with unlabelled Fmoc-Cl under the operating conditions formed (17:83 *FumH<sub>2</sub>*:*FumD<sub>2</sub>*)-4, with a loss of D<sub>2</sub>-label from 63% to 10% incorporation after 18 h, as shown by mass spectrometry. Switching the chemical



**Figure 1 | Operation of a chemically fuelled [2]catenane rotary motor.**

The benzylic amide macrocycle (blue) binds to one or other of the two fumaramide sites (green) of the cyclic track. Bulky groups (red) sterically block passage of the small blue ring and trap it in one compartment or the other (the right- or left-hand side of the track as shown). Cleavage of one of the bulky groups through a chemical reaction (loss of orange ball) allows the small ring to shuttle back and forth between the two fumaramide sites on the track via Brownian motion along the unblocked pathway. Attachment of another bulky group (addition of red ball) through another chemical reaction (under the same conditions) locks in any change of location of the small ring (that is, if the ring has changed compartment it is prevented from returning to the original one). If the kinetics for blocking group attachment are faster when the small ring is far from the reactive site ( $k_{\text{far-attach}} > k_{\text{close-attach}}$ ; for example, for steric reasons), but the cleavage reaction occurs at a rate independent of the small ring position ( $k_{\text{far-cleave}} = k_{\text{close-cleave}}$ ), then the small ring will directionally rotate around the larger one. One of the fumaramide groups is deuterium-labelled to distinguish the compartments and allow the location of the small ring to be determined by <sup>1</sup>H NMR spectroscopy. Compound 1 is the catenane with two Fmoc groups attached. Compound 2 is the catenane with one Fmoc group attached close to the labelled fumaramide group. Compound 2' is the catenane with one Fmoc group attached close to the unlabelled fumaramide group. The italicised prefix (*FumH<sub>2</sub>*- or *FumD<sub>2</sub>*-) refers to the location of the benzylic amide macrocycle in 1, 2 or 2'. Thick arrows indicate the major pathway of a reaction, dashed arrows indicate the minor pathway and thin arrows indicate pathways that occur at similar rates. The blue arrow indicates the direction of net transport of the benzylic amide macrocycle when  $k_{\text{far-attach}} > k_{\text{close-attach}}$  and  $k_{\text{far-cleave}} = k_{\text{close-cleave}}$ .



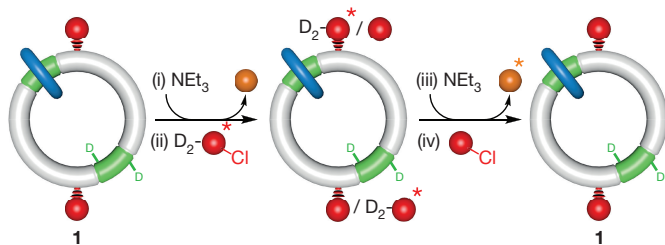
**Figure 2 | [2]Rotaxane model system to demonstrate directional bias for Fmoc addition and position-independent Fmoc cleavage.** **a**, Positional bias of the macrocycle in Fmoc attachment to rotaxane 3. Reaction conditions: Fmoc-Cl (5 equivalents), (R)-5 (5 equivalents), CH<sub>2</sub>Cl<sub>2</sub>, room temperature, duration 18 h. **b**, Partial <sup>1</sup>H NMR spectra (600 MHz, CD<sub>3</sub>OD:CDCl<sub>3</sub> 3:1, 300 K) of 100:0 *FumH*<sub>2</sub>-4:*FumD*<sub>2</sub>-4 (obtained from an unambiguous synthetic route) (top <sup>1</sup>H NMR spectrum) and 17:83 *FumH*<sub>2</sub>-4:*FumD*<sub>2</sub>-4 formed using (R)-5 as the carbonate-forming catalyst (bottom <sup>1</sup>H NMR spectrum). Residual solvent peaks are shown in grey. The lettering corresponds to the proton labelling in **a**. Full spectral assignments are given in Supplementary Information, section 1.3.

fuel being added to the deuterium-labelled version (D<sub>2</sub>-Fmoc-Cl) then fully restored the labelled form of D<sub>2</sub>-(17:83 *FumH*<sub>2</sub>:*FumD*<sub>2</sub>)-4 after 66 h (see Supplementary Fig. 4).

We note that these results indicate that a macrocycle on a polymer consisting of repeat units of 4 without the terminal stopper groups should inexorably be transported towards one end of the polymer chain by treatment with the Fmoc-Cl fuel under these reaction conditions. In other words, rotaxane 4 is a functioning engine system for a chemically fuelled linear molecular motor.

We applied the same principles to the synthesis and operation of a chemically fuelled [2]catenane rotary molecular motor. [2]Catenane 2' was prepared from (R)-3-amino-1,2-propanediol (see Supplementary Information, sections 1.2.2 and 1.3.2). The benzylic amide macrocycle distribution between the two fumaramide sites in 2' is approximately 40:60 *FumH*<sub>2</sub>:*FumD*<sub>2</sub> (estimated from the <sup>1</sup>H NMR shielding of the *FumH*<sub>2</sub> protons in 2' compared to *FumH*<sub>2</sub>-1 and consistent with the

results of carbonate formation promoted by pyridine, a small catalyst (Supplementary Table 2)). The energy barrier for macrocycle exchange between the fumaramide sites in related rotaxanes is ~16 kcal mol<sup>-1</sup> in CDCl<sub>3</sub> (ref. 25), suggesting that macrocycle shuttling in 2' occurs hundreds of times a second under the motor operating conditions. The Fmoc attachment-cleavage chemistry of 2 (Supplementary Information, section 3) mirrored that of rotaxane 3. When catenane 2' was treated with Fmoc-Cl in the presence of catalyst (R)-5, the macrocycle was predominantly trapped in the *FumH*<sub>2</sub> compartment (80:20 *FumH*<sub>2</sub>-1:*FumD*<sub>2</sub>-1, Supplementary Fig. 5a and b), that is,  $k_{\text{far-attach}} > k_{\text{close-attach}}$ . A model catenane was also prepared (see Supplementary Fig. 5c and d), replacing one Fmoc carbonate group of 1 with an analogous Fmoc-methinyl ester. The substitution of an oxygen atom for a carbon atom ensures that cleavage of this group cannot occur under the motor operating conditions. This enabled the Fmoc cleavage reaction to be studied in a catenane possessing only one



Composition				Percentage of D <sub>2</sub> -labelled Fmoc incorporated	
Cycles	Time (h)	1	2/2'	1	D <sub>2</sub> -1
1	0	100	0	100	0
	(i) 6	32	57	–	–
	(ii) 24	94	6	39	61
2	(iii) 29	34	55	–	–
	(iv) 32	100	0	68	32

**Figure 3 | Exchange of Fmoc groups during stepwise operation of catenane 1.** Reaction conditions: (i) NEt<sub>3</sub> (8 equivalents), (ii) D<sub>2</sub>-Fmoc-Cl (16 equivalents), (R)-5 (16 equivalents), (iii) NEt<sub>3</sub> (10 equivalents), (iv) Fmoc-Cl (16 equivalents), (R)-5 (16 equivalents). Some diol (up to 11%) is also formed during the Fmoc-cleavage steps under these conditions (10 equivalents NEt<sub>3</sub>), see Supplementary Table 1. Asterisks indicate D-labelled Fmoc (red ball) or dibenzofulvene (orange ball). The percentage of the D<sub>2</sub>-labelled Fmoc incorporated into the catenane was determined by relative intensity of the [M + Na]<sup>+</sup> signal in electrospray mass spectrometry. Dashes indicate that the degree of deuterium incorporation was not measured at this point.

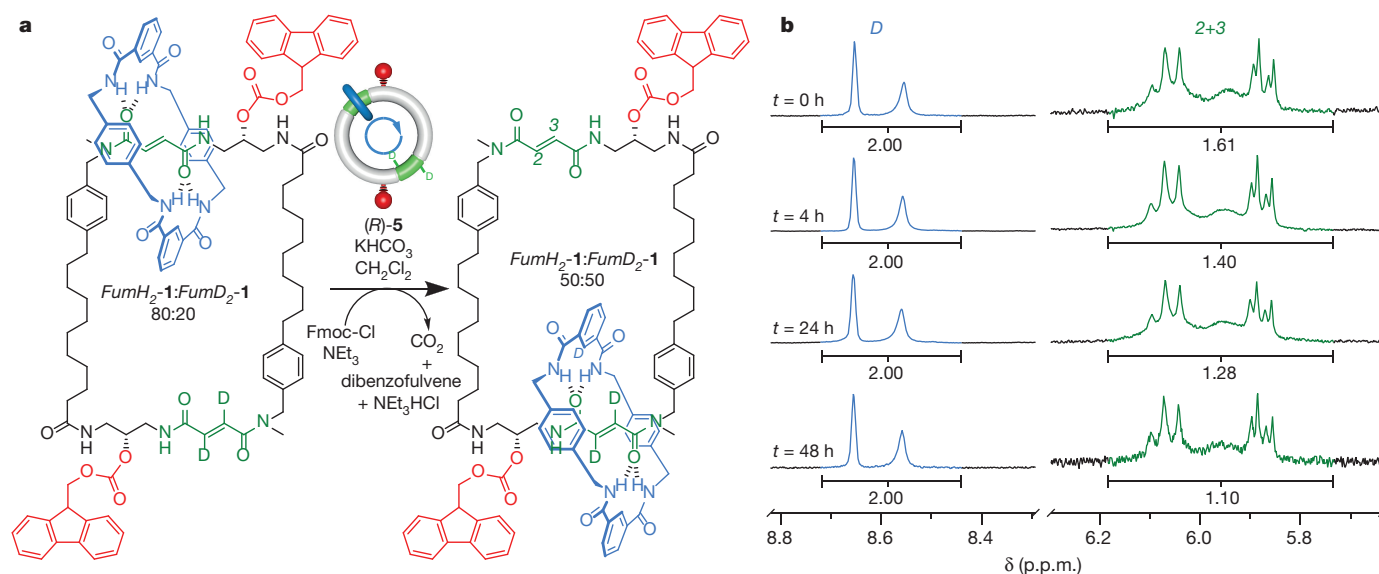
detachable group. As with the rotaxane model, the control experiment demonstrated that the rate of Fmoc cleavage is not affected by the position of the macrocycle in the catenane (Supplementary Fig. 5d), that is,  $k_{\text{far-cleave}} = k_{\text{close-cleave}}$ .

To confirm that both the Fmoc addition and cleavage reactions take place with catenane 1, we first demonstrated that the reactions can occur sequentially (Fig. 3). Catenane 1 was treated with NEt<sub>3</sub> (8 equivalents) and after 6 h 57% of the catenanes had lost one Fmoc group

(forming 2/2'; Fig. 3, cycle 1) with a further 11% of catenanes having had both Fmoc groups cleaved. At this point deuterium-labelled fuel, D<sub>2</sub>-Fmoc-Cl, activated with (R)-5 was added, leading to almost complete derivatization of the catenane hydroxyl groups after 24 h (94% 1). At the end of this cleavage-addition cycle (Fig. 3, cycle 1) electrospray mass spectrometry confirmed that the D<sub>2</sub>-labelled Fmoc groups had been incorporated from the fuel into the catenane motor (Fig. 3, cycle 1). The resulting mixture was then treated with a second cycle of NEt<sub>3</sub>, leading after 5 h to 55% of the catenane with only one Fmoc group, 2/2' (Fig. 3, cycle 2). Subsequent addition of unlabelled Fmoc-Cl regenerated 1 with a majority of the Fmoc groups without deuterium labels (Fig. 3, cycle 2). Thus, over two complete operational cycles, the catenane molecules are shown to sequentially cleave and then add an Fmoc group from the fuel being supplied during that cycle, then cleave and add another Fmoc group from a second batch of fuel.

To monitor the catenane rotary motor during autonomous operation, catenane 1 with 80% of the small rings on the unlabelled fumaramide binding site (80:20 *FumH<sub>2</sub>:FumD<sub>2</sub>*-1) was treated with Fmoc-Cl, (R)-5, Et<sub>3</sub>N and KHCO<sub>3</sub> in CH<sub>2</sub>Cl<sub>2</sub> (Fig. 4). For autonomous operation we used conditions under which the Fmoc groups are added by the Fmoc-Cl fuel and cleaved with no discernible accumulation of diol (that is, 1.5 equivalents Et<sub>3</sub>N instead of the 8 to 10 equivalents employed in the sequential operations) and the distribution of the catenane positional isomers was measured over time by <sup>1</sup>H NMR spectroscopy (Fig. 4). Under these conditions the initial macrocycle distribution changed from 80:20 *FumH<sub>2</sub>:FumD<sub>2</sub>* to 55:45 (Fig. 4b). Shortly after the supply of Fmoc-Cl fuel is cut off, no further change in the distribution of the rings between the compartments occurs (that is, the motor stops working). However, cleavage of the Fmoc groups slowly continues, unless the basic reaction medium is quenched, forming 2/2' and eventually catenane diol.

The ratio of the distribution of the rings between the compartments falls towards 1:1 as a direct consequence of the functioning of the motor as each Fmoc-cleavage reaction serves to equilibrate the distribution of rings between the compartments. Although the Fmoc attachment reaction biases clockwise rotation of the small ring around the track, it does not bias its average position on the track. This is because Fmoc attachment to catenane 2' biases the small ring to the left-hand compartment, whereas Fmoc attachment to catenane 2 biases the small ring to the right-hand compartment. As was demonstrated for



**Figure 4 | Directional transport of the macrocycle monitored by <sup>1</sup>H NMR spectroscopy.** a, Reaction conditions: (R)-5 (5 equivalents), KHCO<sub>3</sub> (20 equivalents), CH<sub>2</sub>Cl<sub>2</sub>, room temperature, Fmoc-Cl, CH<sub>2</sub>Cl<sub>2</sub>, added via syringe pump at 2.4 equivalents per hour, then NEt<sub>3</sub> (1.5 equivalents) after 1 h of Fmoc-Cl addition. b, Partial <sup>1</sup>H NMR spectra

(500 MHz, CD<sub>2</sub>Cl<sub>2</sub>:CD<sub>3</sub>OD 1:1, 300 K) of 80:20 *FumH<sub>2</sub>*-1:*FumD<sub>2</sub>*-1 and after operation for 4 h, 24 h and 48 h. The region 5.7–6.3 p.p.m. is scaled vertically 6× compared to region 8.3–8.8 p.p.m. The two macrocycle positional isomers of catenane 1 each exist as four tertiary amide rotamers.



rotaxane **4** (Fig. 2), the change in the macrocycle distribution that occurs with catenane **1** under the operating conditions (Fig. 4) shows that the Fmoc groups are being cleaved, allowing the small ring to move between compartments by Brownian motion, and the transiently generated hydroxyl groups are being derivatized under consumption of the Fmoc-Cl fuel, that is, confirming that the catenane rotary motor operates autonomously as long as unspent chemical fuel is present.

Proving directional rotation in molecular motors is challenging, not least because each rotational cycle returns the motor components to their starting positions. Evidence for directionality in stepwise-operated small-molecule motors has previously been provided by determining the position of the components at multiple points in a motor's cycle and determining the rates of different pathways to those positions<sup>12,25</sup>. However, in a continuously operating motor with only two minimum energy positions of the components, such as **1**, this approach is not possible. Nevertheless, fuel-driven directional rotation in **1** could be unequivocally established through a series of individually provable premises, a form of deductive logic commonly used in mathematical proofs. If all of the premises are experimentally demonstrated to be correct, and the terms linking the premises to the conclusion are valid, then the conclusion reached is necessarily true. Net directional rotation of the chemically fuelled rotary motor **1** through an information ratchet mechanism (see Supplementary Information, section 6, for an explanation of how directionality intrinsically follows from the rate equations) is demonstrated through experimental verification of each of three premises:

(1) Under the motor operating conditions the Fmoc attachment reaction to the catenane (from Fmoc-Cl) and the Fmoc cleavage reaction from the catenane (forming CO<sub>2</sub> and dibenzofulvene) both occur. This is shown by the experiments that demonstrate that D<sub>2</sub>-Fmoc groups add to the catenane when using D<sub>2</sub>-Fmoc fuel and are then replaced by unlabelled Fmoc groups upon switching to unlabelled fuel (Fig. 3). Fmoc-Cl is not simply being destroyed in the reaction, it is being continuously added and cleaved from the catenane under the operating conditions.

(2) Under the motor operating conditions Fmoc attachment to the catenane hydroxyl group (2 or 2') results in a bias in the distribution of the macrocycle between the compartments in the resulting di-Fmoc catenane (**1**) (that is,  $k_{\text{far-attach}} \neq k_{\text{close-attach}}$ ). This is shown by the positional bias in the Fmoc attachment to 2' experiments (Supplementary Fig. 5a and b), analogous to that shown for rotaxane **3** in Fig. 2a.

(3) Under the motor operating conditions cleavage of one Fmoc group from catenane **1** occurs at a rate independent of the position of the macrocycle (that is,  $k_{\text{far-cleave}} = k_{\text{close-cleave}}$ ). This is proved by the catenane Fmoc cleavage experiments (Supplementary Fig. 5c and d), analogous to that shown for rotaxane 20:80 FumH<sub>2</sub>:FumD<sub>2</sub>-**4** in Fig. 2c.

The effects of the net-directional movement of the rings around the catenane track are directly observed in the experiment shown in Fig. 4. The catenane ring distribution can only change through the benzylic amide macrocycles shuttling between the fumaramide sites when an Fmoc group is transiently cleaved, and the directional bias of the ring movement under these conditions is that experimentally determined in proving premises (2) and (3).

Chirality is not necessary for directional rotation: the wheels of a bicycle travelling down a road rotate clockwise with respect to an observer of one side of the road and counter-clockwise with respect to an observer on the other<sup>25</sup>. However, the chiral centres of **1** differentiate the two faces of the track, defining the direction of the ring rotation in **1** as clockwise with respect to the (R,R)-stereochemistry of the molecular motor.

Just as motor proteins are catalysts for the hydrolysis of ATP, the catenane (**1**) and rotaxane (**4**) motors are catalysts for the conversion of Fmoc-Cl and Et<sub>3</sub>N into dibenzofulvene, CO<sub>2</sub> and Et<sub>3</sub>NHCl. For both the biological and synthetic motors it is the free energy released by the motor-catalysed exergonic reactions that drives the directional displacement of the motor components. In principle, the consumption

of two molecules of Fmoc-Cl is required to power one 360° clockwise ratcheted rotation of the benzylic amide macrocycle around the catenane track of **1**. In practice, the directionality of the fuelled rotation is good: the 80:20 positional bias observed in the Fmoc-attachment reaction means that for every ten molecules of Fmoc-Cl that react with the track the benzylic amide macrocycle makes on average three net directional full rotations about the track. However, unlike motor proteins<sup>20</sup>, rotaxane **3** and catenane **2/2'** are poor catalysts for the destruction of their chemical fuel, and **1** and **4** react with base to form CO<sub>2</sub> and dibenzofulvene only a few times faster than the background base-promoted decomposition of Fmoc-Cl.

From the rate at which the ratio of the macrocycle distribution between the compartments in the catenane falls to unity, the speed of net-directional rotation in the experiment shown in Fig. 4b can be calculated to be ~12 h for each 360° rotation. This might be increased by raising the temperature and/or increasing the concentration and/or rate of addition of the fuel, but changes in these parameters might also affect the net directionality of rotation.

Synthetic chemically fuelled molecular motors **1** and **4** join light-driven molecular rotary motors as engines with the potential to power tasks in molecular nanotechnology<sup>19</sup>. Finding ways to link the position of the ring to more effective catalytic decomposition of the fuel should allow for the development of faster and more efficient small-molecule motors powered by chemical fuels.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 19 January; accepted 12 April 2016.**

- Schliwa, M. & Woehlke, G. Molecular motors. *Nature* **422**, 759–765 (2003).
- Erbas-Cakmak, S., Leigh, D. A., McTernan, C. T. & Nussbaumer, A. L. Artificial molecular machines. *Chem. Rev.* **115**, 10081–10206 (2015).
- Jiménez, M. C., Dietrich-Buchecker, C. & Sauvage, J.-P. Towards synthetic molecular muscles: contraction and stretching of a linear rotaxane dimer. *Angew. Chem. Int. Ed.* **39**, 3284–3287 (2000).
- Bruns, C. J. & Stoddart, J. F. Rotaxane-based molecular muscles. *Acc. Chem. Res.* **47**, 2186–2199 (2014).
- Thordarson, P., Bijsterveld, E. J. A., Rowan, A. E. & Nolte, R. J. M. Epoxidation of polybutadiene by a topologically linked catalyst. *Nature* **424**, 915–918 (2003).
- Lewandowski, B. et al. Sequence-specific peptide synthesis by an artificial small-molecule machine. *Science* **339**, 189–193 (2010).
- Serrelli, V., Lee, C.-F., Kay, E. R. & Leigh, D. A. A molecular information ratchet. *Nature* **445**, 523–527 (2007).
- Ragazzon, G., Baroncini, M., Silvi, S., Venturi, M. & Credi, A. Light-powered autonomous and directional molecular motion of a dissipative self-assembling system. *Nature Nanotechnol.* **10**, 70–75 (2014).
- Cheng, C. et al. An artificial molecular pump. *Nature Nanotechnol.* **10**, 547–553 (2015).
- von Delius, M., Geertsema, E. M. & Leigh, D. A. A synthetic small molecule that can walk down a track. *Nature Chem.* **2**, 96–101 (2010).
- Kassem, S., Lee, A. T. L., Leigh, D. A., Markevicius, A. & Solà, J. Pick-up, transport and release of a molecular cargo using a small-molecule robotic arm. *Nature Chem.* **8**, 138–143 (2016).
- Koumura, N., Zijlstra, R. W. J., van Delden, R. A., Harada, N. & Feringa, B. L. Light-driven monodirectional molecular rotor. *Nature* **401**, 152–155 (1999).
- Eelkema, R. et al. Molecular machines: nanomotor rotates microscale objects. *Nature* **440**, 163 (2006).
- Greb, L. & Lehn, J.-M. Light-driven molecular motors: imines as four-step or two-step unidirectional rotors. *J. Am. Chem. Soc.* **136**, 13114–13117 (2014).
- Li, Q. et al. Macroscopic contraction of a gel induced by the integrated motion of light-driven molecular motors. *Nature Nanotechnol.* **10**, 161–165 (2015).
- Greb, L., Eichhöfer, A. & Lehn, J.-M. Synthetic molecular motors: thermal N inversion and directional photoinduced C=N bond rotation of camphorquinone imines. *Angew. Chem. Int. Ed.* **54**, 14345–14348 (2015).
- Tierney, H. L. et al. Experimental demonstration of a single-molecule electric motor. *Nature Nanotechnol.* **6**, 625–629 (2011).
- Perera, U. G. E. et al. Controlled clockwise and anticlockwise rotational switching of a molecular motor. *Nature Nanotechnol.* **8**, 46–51 (2012).
- Kay, E. R. & Leigh, D. A. Rise of the molecular machines. *Angew. Chem. Int. Ed.* **54**, 10080–10088 (2015).
- Astumian, R. D. Microscopic reversibility as the organizing principle of molecular machines. *Nature Nanotechnol.* **7**, 684–688 (2012).
- Feynman, R. P., Leighton, R. B. & Sands, M. *The Feynman Lectures on Physics* Vol. 1, Ch. 46 (Addison-Wesley, 1963).

22. Kelly, T. R., Tellitu, I. & Sestelo, J. P. In search of molecular ratchets. *Angew. Chem. Int. Edn Engl.* **36**, 1866–1868 (1997).
23. Kelly, T. R., De Silva, H. & Silva, R. A. Unidirectional rotary motion in a molecular system. *Nature* **401**, 150–152 (1999).
24. Kelly, T. R. *et al.* Progress toward a rationally designed, chemically powered rotary molecular motor. *J. Am. Chem. Soc.* **129**, 376–386 (2007).
25. Leigh, D. A., Wong, J. K. Y., Dehez, F. & Zerbetto, F. Unidirectional rotation in a mechanically interlocked molecular rotor. *Nature* **424**, 174–179 (2003).
26. Hernández, J. V., Kay, E. R. & Leigh, D. A. A reversible synthetic rotary molecular motor. *Science* **306**, 1532–1537 (2004).
27. Fletcher, S. P., Dumur, F., Pollard, M. M. & Feringa, B. L. A reversible, unidirectional molecular rotary motor driven by chemical energy. *Science* **310**, 80–82 (2005).
28. Haberhauer, G. A molecular four-stroke motor. *Angew. Chem. Int. Ed.* **50**, 6415–6418 (2011).
29. Lu, C.-H., Ceconello, A., Elbaz, J., Credi, A. & Willner, I. A three-station DNA catenane rotary motor with controlled directionality. *Nano Lett.* **13**, 2303–2308 (2013).
30. Davis, A. P. Tilting at windmills? The second law survives. *Angew. Chem. Int. Ed.* **37**, 909–910 (1998).
31. Alvarez-Pérez, M., Goldup, S. M., Leigh, D. A. & Slawin, A. M. Z. A chemically-driven molecular information ratchet. *J. Am. Chem. Soc.* **130**, 1836–1838 (2008).
32. Carlone, A., Goldup, S. M., Lebrasseur, N., Leigh, D. A. & Wilson, A. A three-compartment chemically-driven molecular information ratchet. *J. Am. Chem. Soc.* **134**, 8321–8323 (2012).
33. Cheng, C., McGonigal, P. R., Stoddart, J. F. & Astumian, R. D. Design and synthesis of nonequilibrium systems. *ACS Nano* **9**, 8672–8688 (2015).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank D. R. Astumian for the analysis of the catenane motor reaction kinetics, the European Research Council (ERC) for funding and the EPSRC National Mass Spectrometry Service Centre (Swansea, UK) for high-resolution mass spectrometry.

**Author Contributions** M.R.W., A.C., J.S., S.M.G. and N.L. carried out the experimental work. M.R.W. and J.S. designed and performed the operation experiments. D.A.L. directed the research. All the authors contributed to the analysis of the results and the writing of the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.A.L. ([david.leigh@manchester.ac.uk](mailto:david.leigh@manchester.ac.uk)).

## METHODS

**General method for autonomous operation of rotary catenane motor 1.** To a solution of **1** (5 mg, 2.6  $\mu\text{mol}$ ) in  $\text{CH}_2\text{Cl}_2$  (0.3 ml) was added (*R*)-**5** (5 equivalents, 8.2 mg, 13.0  $\mu\text{mol}$ ) and  $\text{KHCO}_3$  (20 equivalents, 5.2 mg, 52  $\mu\text{mol}$ ). A solution of Fmoc-Cl (240 mg, 0.93 mmol) in  $\text{CH}_2\text{Cl}_2$  (1.0 ml) was added at a rate of 6.7  $\mu\text{l h}^{-1}$ . After 1 h  $\text{NEt}_3$  (1.5 equivalents, 0.55  $\mu\text{l}$ , 3.9  $\mu\text{mol}$ ) was added and Fmoc-Cl addition continued at a rate of 6.7  $\mu\text{l h}^{-1}$  (6.2  $\mu\text{mol h}^{-1}$  Fmoc-Cl for 2.6  $\mu\text{mol}$  of **1**) for as long as the motor was required to run. After full consumption of the chemical fuel (Fmoc-Cl) the catenane motor was recovered by addition of 1 M HCl (aqueous) (10 ml) and the aqueous layer extracted with  $\text{CH}_2\text{Cl}_2$  ( $3 \times 20$  ml). The combined organic layers were washed with brine, dried over  $\text{Na}_2\text{SO}_4$  and concentrated under reduced pressure. Purification by preparative thin layer chromatography ( $\text{SiO}_2$ ,  $\text{CH}_2\text{Cl}_2$ :EtOH 95:5) gave pristine **1**.



# FeO<sub>2</sub> and FeOOH under deep lower-mantle conditions and Earth's oxygen–hydrogen cycles

Qingyang Hu<sup>1,2\*</sup>, Duck Young Kim<sup>1,2\*</sup>, Wenge Yang<sup>1,3\*</sup>, Liuxiang Yang<sup>1,3</sup>, Yue Meng<sup>4</sup>, Li Zhang<sup>1,2</sup> & Ho-Kwang Mao<sup>1,2</sup>

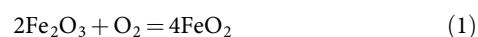
The distribution, accumulation and circulation of oxygen and hydrogen in Earth's interior dictate the geochemical evolution of the hydrosphere, atmosphere and biosphere<sup>1</sup>. The oxygen-rich atmosphere and iron-rich core represent two end-members of the oxygen–iron (O–Fe) system, overlapping with the entire pressure–temperature–composition range of the planet. The extreme pressure and temperature conditions of the deep interior alter the oxidation states<sup>1</sup>, spin states<sup>2</sup> and phase stabilities<sup>3,4</sup> of iron oxides, creating new stoichiometries, such as Fe<sub>4</sub>O<sub>5</sub> (ref. 5) and Fe<sub>5</sub>O<sub>6</sub> (ref. 6). Such interactions between O and Fe dictate Earth's formation, the separation of the core and mantle, and the evolution of the atmosphere. Iron, in its multiple oxidation states, controls the oxygen fugacity and oxygen budget, with hydrogen having a key role in the reaction of Fe and O (causing iron to rust in humid air). Here we use first-principles calculations and experiments to identify a highly stable, pyrite-structured iron oxide (FeO<sub>2</sub>) at 76 gigapascals and 1,800 kelvin that holds an excessive amount of oxygen. We show that the mineral goethite, FeOOH, which exists ubiquitously as 'rust' and is concentrated in bog iron ore, decomposes under the deep lower-mantle conditions to form FeO<sub>2</sub> and release H<sub>2</sub>. The reaction could cause accumulation of the heavy FeO<sub>2</sub>-bearing patches in the deep lower mantle, upward migration of hydrogen, and separation of the oxygen and hydrogen cycles. This process provides an alternative interpretation for the origin of seismic and geochemical anomalies in the deep lower mantle, as well as a sporadic O<sub>2</sub> source for the Great Oxidation Event over two billion years ago that created the present oxygen-rich atmosphere.

We started with α-Fe<sub>2</sub>O<sub>3</sub> (haematite) powder loaded in cryogenically condensed liquid O<sub>2</sub> in the sample chamber of a diamond-anvil cell (DAC) (see Methods). The pressure was initially raised to 78 GPa; no reaction between haematite and O<sub>2</sub> was observed at ambient temperature (~293 K). Using a Nd-doped Y<sub>3</sub>Al<sub>5</sub>O<sub>12</sub> laser system<sup>7</sup> to heat the sample to 1,800 K *in situ* at high pressure, the sample became semi-transparent (Fig. 1b), suggesting that a chemical reaction had occurred. The X-ray diffraction (XRD) pattern shows new sets of sharp, single crystal-like diffraction spots (Fig. 1a) that are readily distinguishable from the original broad and smooth texture of the Fe<sub>2</sub>O<sub>3</sub> powder pattern. Integration of the diffraction spots (Fig. 2a) shows eight peaks that do not match any known Fe<sub>2</sub>O<sub>3</sub> (refs 3, 4) or O<sub>2</sub> phases<sup>8</sup>, but can be unambiguously indexed to a rather simple cubic structure (Fig. 1c) with the space group *Pa*3̄ (Table 1).

The spotty XRD pattern is ideally suited to the multigrain crystallography method<sup>9</sup> recently adopted for high-pressure research<sup>10,11</sup>. The spots are treated as diffraction from multiple single crystals, and sorted according to individual crystal orientation matrices. At least 33 single crystallites were identified by the multigrain crystallography method software. All symmetry-allowed spots for *Pa*3̄ are present, and all observed spots can be accounted for by the *Pa*3̄ unit cell. The details of five crystallites are presented in Extended Data Tables 1 and 2.

The new phase has a structure identical to that of pyrite (FeS<sub>2</sub>) with oxygen replacing sulphur, the next-row chalcogen element. Results from Rietveld refinement are shown in Fig. 2. For this structure, oxygen atoms not only form O–Fe bonds of 1.792 Å, but also O–O bonds of 1.937 Å (Extended Data Fig. 1 and Extended Data Table 3), that are typical of peroxide. Analogous to the archetypical pyrite, the iron in FeO<sub>2</sub> is considered to be ferrous. Curiously, the oxidation of Fe<sub>2</sub>O<sub>3</sub> to FeO<sub>2</sub> reduces Fe<sup>3+</sup> to Fe<sup>2+</sup>. This can be understood with the concurrent oxidation of O<sup>2–</sup> to O<sup>2–</sup> and O<sup>0</sup> as indicated by the O–O bond. In other words, this material can be viewed as FeO holding extra O<sub>2</sub>. We shall refer to the pyrite phase of *Pa*3̄ peroxide as the P-phase.

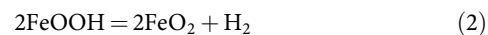
To assess the stability of the P-phase under pressure, we calculated the volume change of the reaction at 76 GPa as follows:



Here we used molar volumes of 35.69 Å<sup>3</sup> for Fe<sub>2</sub>O<sub>3</sub> in the *Aba*2 structure<sup>4</sup>, 12.79 Å<sup>3</sup> for O<sub>2</sub> in the O<sub>8</sub> cluster<sup>12</sup>, and 20.76 Å<sup>3</sup> for the P-phase. The reaction has a volume shrink of ΔV/V = –1.4%. Pressure lowers the Gibbs free energy of the reaction by ΔG = ∫ΔVdP, thus favouring the formation of FeO<sub>2</sub> at increasing pressure. The P-phase is non-quenchable to ambient conditions; its XRD peaks disappear below 31 GPa during pressure release at 300 K (Extended Data Fig. 2).

We expanded our study from the O–Fe binary to the O–Fe–H ternary, and showed that the P-phase could also be synthesized under moderately reducing conditions coexisting with H<sub>2</sub>. We studied the Fe<sub>2</sub>O<sub>3</sub>–H<sub>2</sub>O join, in which the most stable compound FeOOH occurs ubiquitously as rust on Earth's surface, in the deep ocean, on meteorites, on other planets, and on moons, in the α-, β-, γ-, δ-, or ε-FeOOH forms. It concentrates in bog iron ore deposits, which have been used as a copious, renewable resource of iron ever since the Iron Age. The α-FeOOH (goethite) transforms to the ε-phase at high pressure and decomposes to Fe<sub>2</sub>O<sub>3</sub> + H<sub>2</sub>O at high temperature. Its pressure–temperature phase boundaries and pressure–volume–temperature equations of state have been previously determined up to 29.4 GPa and 523 K (ref. 13).

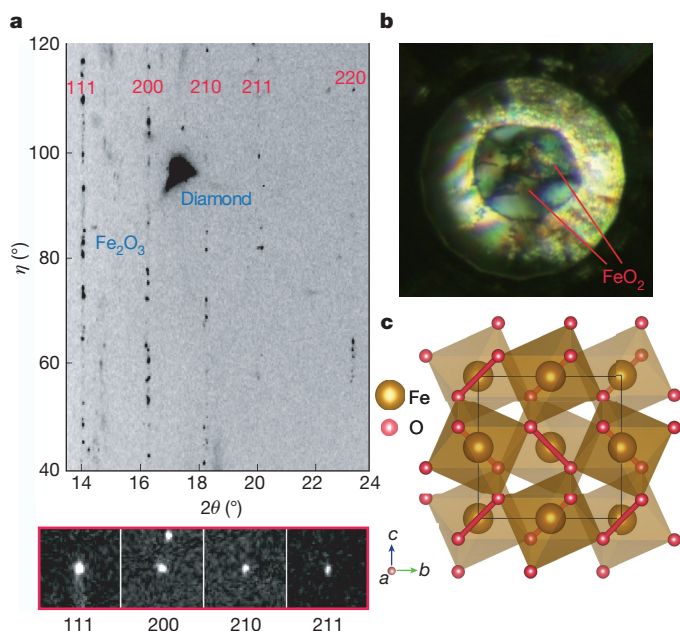
We compressed goethite in Ne pressure medium to 92 GPa and laser-heated it to 2,050 K. XRD clearly shows the conversion to the P-phase (Fig. 2b), indicating the following reaction:



The H<sub>2</sub> at 92 GPa and 2,050 K is far above its melting temperature of 900 K (ref. 14), and the H<sub>2</sub> fluid is highly mobile. Raman spectroscopy is used to search for H<sub>2</sub>, and clearly observed H<sub>2</sub> vibron peaks at 5,180 cm<sup>–1</sup> (Fig. 3), corresponding to H<sub>2</sub> in the Ne pressure medium<sup>15</sup>. The production of free H<sub>2</sub> indicates a moderately reducing condition.

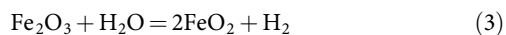
<sup>1</sup>Center for High Pressure Science and Technology Advanced Research (HPSTAR), Shanghai 201203, China. <sup>2</sup>Geophysical Laboratory, Carnegie Institution, Washington DC 20015, USA. <sup>3</sup>High Pressure Synergetic Consortium (HPSynC), Geophysical Laboratory, Carnegie Institution, Argonne, Illinois 60439, USA. <sup>4</sup>High Pressure Collaborative Access Team (HPCAT), Geophysical Laboratory, Carnegie Institution, Argonne, Illinois 60439, USA.

\*These authors contributed equally to this work.



**Figure 1 | The FeO<sub>2</sub> phase.** **a**, A two-dimensional XRD image of the Fe<sub>2</sub>O<sub>3</sub> + O<sub>2</sub> experiment at 76 GPa after laser heating, collected at  $\omega = -8.5^\circ$ , with X-ray wavelength of 0.6199 Å. The original image in  $2\theta$ – $\eta$  polar coordination is converted into Cartesian coordinates. The newly developed sharp spots are from several single crystals of FeO<sub>2</sub>, while smeared powder rings are from the remaining Fe<sub>2</sub>O<sub>3</sub>. Four selected diffraction spots and their Miller indices from the P-phase are shown below. Their observed Bragg angles ( $2\theta$ ), rotation angles ( $\omega$ ) and azimuthal angles ( $\eta$ ) are listed in Extended Data Table 1. **b**, A microphotographic image of FeO<sub>2</sub> through diamond culets. **c**, Structural representation of the pyrite-type FeO<sub>2</sub>.

As a chemical equivalent of Fe<sub>2</sub>O<sub>3</sub> + H<sub>2</sub>O, the goethite reaction (2) demonstrates that the water in deep Earth<sup>16</sup> can provide an abundant source of O<sub>2</sub>, that is:

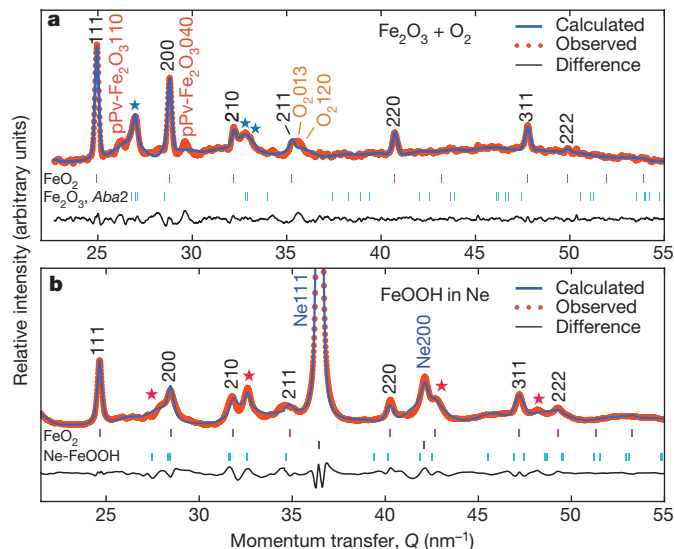


Although the exact quantity of H<sub>2</sub>O in the mantle is uncertain, the existence of H<sub>2</sub>O there in hydrates or other forms is well accepted<sup>17–19</sup>.

We were initially motivated by our computational predictions. A first-principles-based structure-searching algorithm<sup>20</sup> allows us to investigate the energy landscape of Fe–O compounds under pressure. At 100 GPa and 300 GPa, we conducted extensive prediction of Fe–O compounds using this model, and FeO<sub>2</sub> appears as one of the most stable phases. In the convex hull curve (Fig. 4a and b), FeO<sub>2</sub> stands out as energetically extremely favourable (the deepest hull). Our phonon calculations of FeO<sub>2</sub> show stable phonons at all pressure (Fig. 4c). The experimentally observed structural parameters of the P-phase FeO<sub>2</sub> agree exactly with the *ab initio* prediction (Table 1). Electron localization function calculations and Bader analysis shed light on the chemical bonding nature of FeO<sub>2</sub>. Valence electrons near anionic O are highly localized and electron localization function minima are located between Fe and O atoms.

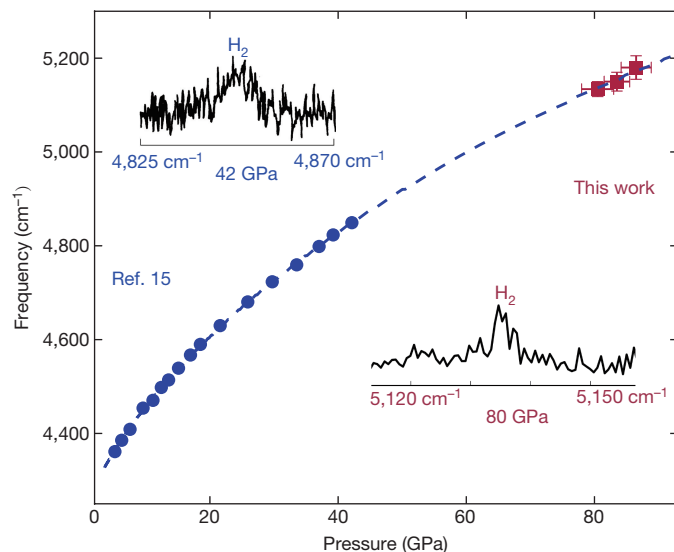
The P-phase has been previously predicted at 100–465 GPa by a computational search for high-pressure iron oxides<sup>21</sup>. It is clearly one of the most prominent phases in the convex hull curve at all of the calculated pressures, which are thought to occur at Earth's centre.

Our experiments and theoretical calculations demonstrate that if the surface assemblage FeOOH or Fe<sub>2</sub>O<sub>3</sub> + H<sub>2</sub>O is thrust deeper than 1,800 km (into the deep lower mantle), it will form the P-phase. The frequent occurrence of such assemblages in the down-going slabs suggests that reaction (2) could have started as early as the accretion of the early Earth from planetesimals of assorted compositions. The water-rich and iron-oxide-rich fragments would release H<sub>2</sub>

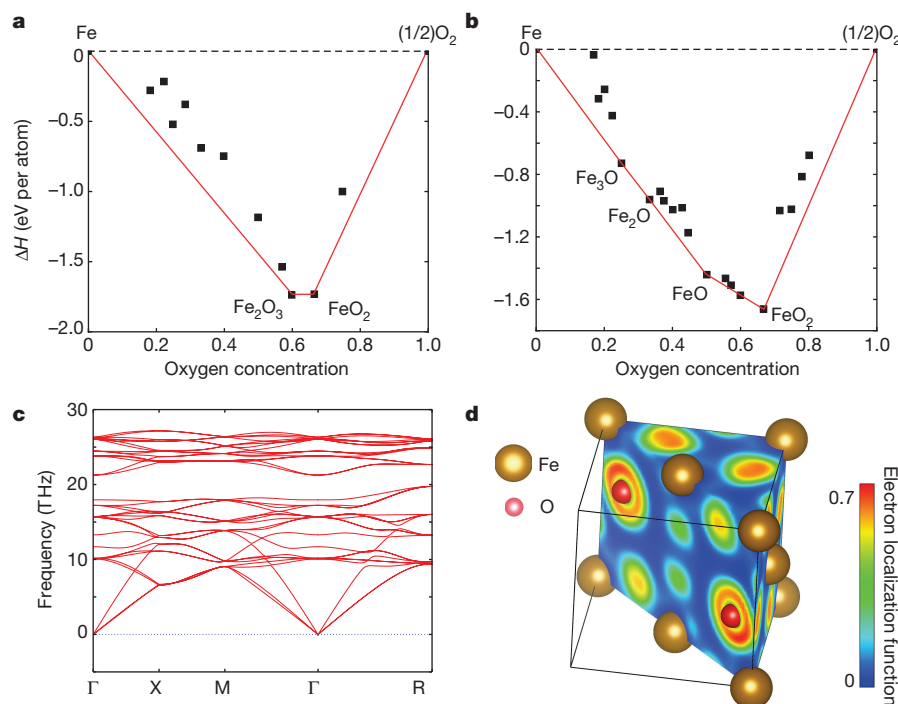


**Figure 2 | Integrated XRD pattern and Rietveld refinement of the pyrite-type FeO<sub>2</sub> phase.** **a**, FeO<sub>2</sub> synthesized from Fe<sub>2</sub>O<sub>3</sub> and O<sub>2</sub> at 76 GPa with  $R$  indices  $R_1 = 0.0663$  and  $wR_2 = 0.171$ . Blue stars belong to the Aba2 phase of Fe<sub>2</sub>O<sub>3</sub>. We attribute the residual peaks to post-perovskite-type (pPv) Fe<sub>2</sub>O<sub>3</sub> and O<sub>2</sub>. **b**, FeO<sub>2</sub> synthesized from FeOOH at 92 GPa, with  $R$  indices,  $R_1 = 0.0541$  and  $wR_2 = 0.167$ . Red stars refer to residual  $\epsilon$ -FeOOH. Similar results are reproduced in multiple experiments for each composition.

and convert to the P-phase when pressure exceeds 76 GPa. With its high density (7.026 g cm<sup>−3</sup> at 76 GPa) in comparison to the density (5 g cm<sup>−3</sup>) of the mantle according to the Preliminary Reference Earth Model<sup>22</sup>, the P-phase would normally settle at depth, while the light and mobile hydrogen would diffuse, infiltrate, or react to form other volatiles, and work its way up to complete the hydrogen cycle. A portion of the H<sub>2</sub> might eventually escape into space. Plate



**Figure 3 | Raman peak of the hydrogen Q<sub>1</sub> vibron in dense neon.** Data points (blue circles) below 45 GPa and the top inset are taken from Loubeyre's hydrogen vibron measurements of H<sub>2</sub> in the Ne matrix<sup>15</sup>. The blue dashed curve is the fourth-order polynomial fitting given by ref. 15. Data points (red squares) above 80 GPa and the bottom inset are taken from the present results of our FeOOH experiment after laser heating and during decompression to 80 GPa. The sharp hydrogen vibron peak at 5,180–5,170 cm<sup>−1</sup> clearly indicates the hydrogen H<sub>2</sub> in the neon matrix. Errors of frequency are calculated from the full-width half-maximum of the Raman peak. Pressure uncertainty is derived from multiple measurements of diamond line shifts at the centre of the culet.



**Figure 4 | Crystal structure search results.** **a**, 100 GPa; **b**, 300 GPa. The convex hull graph referenced by a dashed line shows each compound's stability with respect to decomposition into individual Fe and O elements. Solid squares show the formation enthalpy of each compound possessing the lowest enthalpy within fixed stoichiometry and chemical formula (red

line is a guide to the eye). **c**, Phonon dispersion relations sampling along high-symmetry points in the Brillouin zone ( $\Gamma$ -X-M- $\Gamma$ -R) at 76 GPa. **d**, The electron localization function distribution modelled in the  $\text{FeO}_2$  crystal lattice.

tectonics would have continued to supply  $\text{FeOOH}$ ,  $\text{Fe}_2\text{O}_3$ , and  $\text{H}_2\text{O}$  to the down-going slabs, thus accumulating P-phase to form patches of oxygen reservoirs in the deep lower mantle and at the same time sustaining the hydrogen cycle. The high mobility and rapid cycling of hydrogen would help to build up substantial patches of the P-phase, possibly detectable as seismic anomalies in the  $D''$  layer and other deep lower-mantle regions.

Occasionally, the P-phase-rich patches could have been swept up by plumes or other large-scale mantle dynamic processes. Once they reached the middle lower mantle at depths of less than 1,500 km, the P-phase would decompose according to the reversal of equation (1) to provide a sporadic source of extra  $\text{O}_2$ .

The idea of such P-phase-rich patches is in harmony with the general concept of an overall reducing lower mantle. First, reaction (2) shows that  $\text{FeO}_2$  can coexist with  $\text{H}_2$ , indicating a moderately reducing condition in spite of its high oxygen content. Second, our knowledge of oxygen fugacity in the deep Earth is based on very limited sampling from the lower mantle (with none from the deep lower mantle). Sampling is chemically selective towards cratonic lithosphere<sup>23</sup>. Considering the geochemical diversity of the mantle<sup>24</sup>, the major

elements can vary greatly in regions of subduction crust<sup>25</sup>, upwelling plumes<sup>26</sup>,  $D''$  layer<sup>27</sup>, and many other small patches. Solid diffusion is extremely inefficient<sup>28</sup>, and cannot eliminate the oxygen inhomogeneity between a P-phase-rich patch and the adjacent rocks kilometres away, even at the mantle temperatures occurring through geological time. The oxygen fugacity in the deep lower mantle is probably as inhomogeneous as that in the crust, which ranges from the oxygen fugacity of the iron-wüstite buffer<sup>29</sup> to that of free  $\text{O}_2$  in the air. The oxygen-rich P-phase could exist in pockets, locally or regionally.

The scenarios of spatial and temporal heterogeneity of oxygen in the deep lower mantle have far-reaching implications. The oxygen-rich patches may lead to new phase assemblages with very different mineralogical, chemical and physical signatures from the nominally bridgmanite-ferropericlase mantle. The new phase assemblages may be responsible for many unexplained seismic and geochemical anomalies in the deep lower mantle and the  $D''$  layer. The new scenario introduces a complex picture of the deep lower mantle that calls for in-depth study of the P-phase.

The Great Oxidation Event marked the permanent rise of the  $\text{O}_2$  level in the atmosphere, which did not previously contain free  $\text{O}_2$ . It is thought to have occurred 2.4–2.1 billion years ago<sup>30</sup>, on the basis of evidence such as the appearance of highly oxidized red soil, the disappearance of easily oxidized  $\text{FeS}_2$  pyrite<sup>31</sup>, and the disappearance of distinctive non-mass-dependent sulphur isotope fractionations<sup>32</sup>. In addition to the proposed biogenic origin of the  $\text{O}_2$ , the emergence of P-phase-bearing patches could provide an extra, eventual, abiotic source of  $\text{O}_2$ . Whether the strong uprising of the P-phase patches 2.4–2.1 billion years ago was an accidental, sporadic event or was triggered by some geodynamic instability would be very interesting to know. Exploration of these hypotheses, however, requires further investigations of the physical, chemical, and mineralogical properties of the P-phase under the deep lower-mantle conditions.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Table 1 | Indexed peaks of the XRD pattern from the P-phase at 76 GPa**

<i>hkl</i>	$d_{\text{obs}}$ (Å)	$d_{\text{cal}}$ (Å)	$\Delta d$ (Å)	$d_{\text{simu}}$ (Å)
111	2.5190(1)	2.5189	−0.0001	2.500
200	2.1817(5)	2.1814	−0.0003	2.165
210	1.9508(3)	1.9511	0.0003	1.937
211	1.7821(15)	1.7811	−0.0010	1.768
220	1.5421(4)	1.5425	0.0004	1.531
311	1.3151(5)	1.3154	0.0003	1.306
222	1.2600(7)	1.2594	−0.0006	1.250
320	1.2093(12)	1.2100	0.0007	1.201

The lattice parameter is calculated to be  $a = 4.3628(1)$  Å. The wavelength of the synchrotron X-ray is 0.4344 Å. The  $d$ -spacings at 76 GPa are as follows:  $d_{\text{obs}}$  are the observed peak positions;  $d_{\text{cal}}$  is calculated from the averaged lattice parameter;  $\Delta d = d_{\text{cal}} - d_{\text{obs}}$ ; and  $d_{\text{simu}}$  is the first-principles simulation. The estimated standard deviations are included in parentheses.



Received 14 January; accepted 12 April 2016.

1. Frost, D. J. & McCammon, C. A. The redox state of Earth's mantle. *Annu. Rev. Earth Planet. Sci.* **36**, 389–420 (2008).
2. Badro, J. *et al.* Iron partitioning in Earth's mantle: toward a deep lower mantle discontinuity. *Science* **300**, 789–791 (2003).
3. Ono, S. & Ohishi, Y. In situ X-ray observation of phase transformation in Fe<sub>2</sub>O<sub>3</sub> at high pressures and high temperatures. *J. Phys. Chem. Solids* **66**, 1714–1720 (2005).
4. Bykova, E. *et al.* Structural complexity of simple Fe<sub>2</sub>O<sub>3</sub> at high pressures and temperatures. *Nature Commun.* **7**, 10661 (2016).
5. Lavina, B. *et al.* Discovery of the recoverable high-pressure iron oxide Fe<sub>4</sub>O<sub>5</sub>. *Proc. Natl Acad. Sci. USA* **108**, 17281–17285 (2011).
6. Lavina, B. & Meng, Y. Unraveling the complexity of iron oxides at high pressure and temperature: synthesis of Fe<sub>3</sub>O<sub>6</sub>. *Sci. Adv.* **1**, e1400260 (2015).
7. Meng, Y., Hrubciak, R., Rod, E., Boehler, R. & Shen, G. New developments in laser-heated diamond anvil cell with in-situ x-ray diffraction at High Pressure Collaborative Access Team. *Rev. Sci. Instrum.* **86**, 072201 (2015).
8. Akahama, Y. & Kawamura, H. High-pressure Raman spectroscopy of solid oxygen. *Phys. Rev. B* **54**, R15602–R15605 (1996).
9. Sørensen, H. O. *et al.* Multigrain crystallography. *Z. Kristallogr.* **227**, 63–78 (2012).
10. Zhang, L. *et al.* Single-crystal structure determination of (Mg,Fe)SiO<sub>3</sub> postperovskite. *Proc. Natl Acad. Sci. USA* **110**, 6292–6295 (2013).
11. Zhang, L. *et al.* Disproportionation of (Mg,Fe)SiO<sub>3</sub> perovskite in Earth's deep lower mantle. *Science* **344**, 877–882 (2014).
12. Fujihisa, H. *et al.* O<sub>8</sub> cluster structure of the epsilon phase of solid oxygen. *Phys. Rev. Lett.* **97**, 085503 (2006).
13. Gleason, A. E., Jeanloz, R. & Kunz, M. Pressure–temperature stability studies of FeOOH using X-ray diffraction. *Am. Mineral.* **93**, 1882–1885 (2008).
14. Howie, R. T., Dalladay-Simpson, P. & Gregoryanz, E. Raman spectroscopy of hot hydrogen above 200 GPa. *Nature Mater.* **14**, 495–499 (2015).
15. Loubeyre, P., Letoullec, R. & Pinceaux, J. P. Properties of H<sub>2</sub> under strong compression in a Ne matrix. *Phys. Rev. Lett.* **67**, 3271–3274 (1991).
16. Jacobsen, S. D. & van der Lee, S. (eds) *Earth's Deep Water Cycle* Vol. 168 (American Geophysical Union, 2006).
17. Nishi, M. *et al.* Stability of hydrous silicate at high pressures and water transport to the deep lower mantle. *Nature Geosci.* **7**, 224–227 (2014).
18. Pamato, M. G. *et al.* Lower-mantle water reservoir implied by the extreme stability of a hydrous aluminosilicate. *Nature Geosci.* **8**, 75–79 (2014).
19. Schmandt, B., Jacobsen, S. D., Becker, T. W., Liu, Z. & Dueker, K. G. Dehydration melting at the top of the lower mantle. *Science* **344**, 1265–1268 (2014).
20. Oganov, A. R., Lyakhov, A. O. & Valle, M. How evolutionary crystal structure prediction works—and why. *Acc. Chem. Res.* **44**, 227–237 (2011).
21. Weerasinghe, G. L., Pickard, C. J. & Needs, R. J. Computational searches for iron oxides at high pressures. *J. Phys. Condens. Matter* **27**, 455501 (2015).
22. Dziewonski, A. & Anderson, D. L. Preliminary reference earth model. *Phys. Earth Planet. Inter.* **25**, 297–356 (1981).
23. Stagno, V., Ojwang, D. O., McCammon, C. A. & Frost, D. J. The oxidation state of the mantle and the extraction of carbon from Earth's interior. *Nature* **493**, 84–88 (2013).
24. Li, M., McNamara, A. K. & Garnero, E. J. Chemical complexity of hotspots caused by cycling oceanic crust through mantle reservoirs. *Nature Geosci.* **7**, 366–370 (2014).
25. Hirose, K., Fei, Y., Ma, Y. & Mao, H. K. The fate of subducted basaltic crust in the Earth's lower mantle. *Nature* **397**, 53–56 (1999).
26. Ren, Z. Y., Ingle, S., Takahashi, E., Hirano, N. & Hirata, T. The chemical structure of the Hawaiian mantle plume. *Nature* **436**, 837–840 (2005).
27. Hutko, A. R., Lay, T., Garnero, E. J. & Revenaugh, J. Seismic detection of folded, subducted lithosphere at the core–mantle boundary. *Nature* **441**, 333–336 (2006).
28. Holzapfel, C., Rubie, D., Frost, C. & Langenhorst, F. Fe–Mg interdiffusion in (Mg,Fe)SiO<sub>3</sub> perovskite and lower mantle reequilibration. *Science* **309**, 1707–1710 (2005).
29. Yang, J. *et al.* Diamonds, native elements and metal alloys from chromitites of the Ray-Iz ophiolite of the Polar Urals. *Gondwana Res.* **27**, 459–485 (2015).
30. Lyons, T. W., Reinhard, C. T. & Planavsky, N. J. The rise of oxygen in Earth's early ocean and atmosphere. *Nature* **506**, 307–315 (2014).
31. Canfield, D. E. The early history of atmospheric oxygen: homage to Robert M. Garrels. *Annu. Rev. Earth Planet. Sci.* **33**, 1–36 (2005).
32. Farquhar, J., Bao, H. & Thiemens, M. Atmospheric influence of Earth's earliest sulfur cycle. *Science* **289**, 756–758 (2000).

**Acknowledgements** We thank Y. Fei for providing haematite powder samples; A. Goncharov for conducting the laser-heating treatment; and D. Z. Zhang, J.-F. Shu, J. Smith, Y. Kono, K. Yang, S. Yan, Z. H. Yu, Y. Yuan, M. Q. Hou and L. Xu for beamline technical support. XRD measurements were performed at the High Pressure Collaborative Access Team (HPCAT 16-IDB and 16-BMD) Advanced Photon Source (APS), Argonne National Laboratory, and the BL15U1 beamline, Shanghai Synchrotron Radiation Facility in China. Part of the experiments was performed at the 13BM-C experimental station of the GeoSoilEnviroCARS facility at the APS. HPCAT operations are supported by the DOE-NNSA under award number DE-NA0001974 and by the DOE-BES under award number DE-FG02-99ER45775, with partial instrumentation funding by the NSF. 13BM-C operation is supported by COMPRES through the Partnership for Extreme Crystallography (PX<sup>2</sup>) project, under NSF Cooperative Agreement EAR 11-57758. APS is supported by the DOE-BES, under contract number DE-AC02-06CH11357. Q.H. and H.-K.M. were supported by NSF grants EAR-1345112 and EAR-1447438. L.Z. was supported by the Foundation of President of China Academy of Engineering Physics (grant no. 201402032) and the National Natural Science Foundation of China (grant no. 41574080). This work was also supported in part by the National Natural Science Foundation of China (grant number U1530402).

**Author Contributions** W.Y., L.Y. and Q.H. carried out the experiment. Y.M. conducted the laser heating for the first synthesis of FeO<sub>2</sub>. W.Y., L.Z., Q.H. and H.-K.M. performed the experimental data analysis. D.Y.K. predicted the FeO<sub>2</sub> structure and performed the computer simulation. H.-K.M. conceived and designed the project and directed the calculations and experiments. Q.H. and H.-K.M. wrote the manuscript. All authors contributed to the discussion of the results and revision of the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.-K.M. ([hmao@gl.ciw.edu](mailto:hmao@gl.ciw.edu)).

## METHODS

**XRD of haematite in O<sub>2</sub>.** Angular dispersive XRD experiments were performed at the 16-BMD and 16-IDB stations of the High-Pressure Collaborative Access Team (HPCAT) and the 13BM-C station of the GeoSoilEnviroCARS, at the Advanced Photon Source, Argonne National Laboratory.

High-purity Fe<sub>2</sub>O<sub>3</sub> powders from Alfa Aesar (MFCD00011008, 99.99% purity) were annealed at 1,296 K for 12 h to eliminate the absorbed water. The dehydrated samples were compressed into patties of size 60  $\mu$ m (length)  $\times$  60  $\mu$ m (width)  $\times$  10  $\mu$ m (thickness), and loaded in a DAC. We use diamond anvils with culets of 200  $\mu$ m to access pressures above 70 GPa. The sample chamber was a 100- $\mu$ m-diameter hole drilled in a pre-compressed rhenium gasket. The DAC was placed in a sealed container immersed in liquid nitrogen. O<sub>2</sub> gas was piped into the container. Liquefied O<sub>2</sub> infused into the sample chamber, and served as both oxidant and pressure medium after the chamber was sealed by compression. The initial pressure after loading O<sub>2</sub> was 8 GPa.

The compression rate from 8 to 78 GPa was as slow as 5 GPa per hour. During compression, the Raman signals from samples reproduce the literature results of Fe<sub>2</sub>O<sub>3</sub> (ref. 33) and solid O<sub>2</sub> (ref. 8). We checked the structure of compressed Fe<sub>2</sub>O<sub>3</sub> at 78 GPa, and verified that the diffraction pattern agreed with the Rh<sub>2</sub>O<sub>3</sub>-type Fe<sub>2</sub>O<sub>3</sub> phase<sup>3,4</sup>. The sample was then heated by a double-sided laser system<sup>7</sup> to promote chemical reaction. The heating temperature reached 1,800 K on both sides of the sample, measured by fitting the black-body radiation curve. After laser heating and quenching to ambient temperature, the sample pressure was equilibrated at 76 GPa.

To confirm the synthesis conditions for FeO<sub>2</sub>, we implemented separate runs at different pressures (56–81 GPa) using the same assemblage. We gradually increase the laser-heating power and monitor the change of XRD pattern. FeO<sub>2</sub> was observed above 75 GPa when the temperature was raised above 1,600 K, which represented the kinetic barrier of the reaction. FeO<sub>2</sub> was not observed at pressures below 72 GPa even at temperatures as high as 2,050 K, thus defining the lower stability limit of FeO<sub>2</sub> (Extended Data Fig. 3a).

Diffraction patterns performed at 16-BMD (HPCAT) with a monochromatic X-ray energy of 40.1 keV were collected on a Mar 345 image plate detector while the DAC was rotated so that the sample-beam angle varied from  $-7^\circ$  to  $7^\circ$ . The P-phase is identified (Fig. 2). We performed additional XRD experiments at 13BM-C with rotation angle  $-13.0^\circ$  to  $12.5^\circ$  and X-ray energy of 28.5 keV. Observation of the P-phase was well reproduced with the same experimental conditions. Pressure was determined by calibrating the derivative shift of the diamond Raman mode in an offline Raman system, confirmed by the vibron mode frequency of  $\epsilon$ -O<sub>2</sub> (ref. 8). The pressure uncertainty is up to  $\pm 3$  GPa, derived from the difference between diamond line shifts and O<sub>2</sub> vibron mode shifts.

**Decompress FeO<sub>2</sub>.** The sample was decompressed to the ambient condition from 76 GPa at a rate of 8 GPa per hour. The P-phase was recognized in diffraction patterns until 41 GPa, where the signature (111) and (200) becomes weak in intensity. The FeO<sub>2</sub> structure is eventually invisible at 31 GPa (Extended Data Fig. 2). Three newly emerged peaks are possibly associated with the low-pressure haematite.

**Multigrain single-crystal XRD.** To identify the new unit cell of an unknown phase, the conventional powder XRD method uses only the Miller indices–Bragg angle relationship ( $hkl-2\theta$ ) and the answer is often non-unique. Like the single-crystal XRD method, the multigrain crystallography method requires all observable Miller indices simultaneously to satisfy the stringent geometrical relation among  $2\theta$ ,  $\omega$  (rotation axis perpendicular to the incident X-ray beam), and  $\eta$  (rotation axis parallel to the incident X-ray beam) within a tight uncertainty range, so the unit-cell assignment is absolutely definitive. In addition, the multigrain crystallography method has the statistical advantage over the single-crystal XRD owing to the number of multiple crystals and the coverage advantage enabling the access of full orientations that can be blocked for a single crystal by the limited DAC opening.

Multigrain single-crystal XRD experiments were implemented at BL15U1 station, the Shanghai Synchrotron Radiation Facility and 13BM-C, GeoSoilEnviroCARS. At Shanghai Synchrotron Radiation Facility, diffraction patterns were collected on a charge-couple device detector with nano-focusing incident beam (beam size of 1.5  $\mu$ m  $\times$  2  $\mu$ m) and beam energy of 20.0 keV. A total number of 50 images were collected by rotating the sample beam angle from  $-12.5^\circ$  to  $12.5^\circ$ , with a scanning step of  $0.5^\circ$ . For experiments performed at 13BM-C, the incident X-ray is 15  $\mu$ m  $\times$  15  $\mu$ m in beam size and 28.5 keV in energy. The sample beam angle is  $-13.0^\circ$  to  $12.5^\circ$ . The diffraction peak indexing and  $d$ -spacing for five selected FeO<sub>2</sub> crystal grains are summarized in the Extended Data Table 1a–e

respectively. A short summary of the multigrain diffraction results is shown in Extended Data Table 2.

**XRD experiment on FeOOH.** For the FeOOH experiment, angular dispersive XRD experiments were performed at the 16-IDB station of HPCAT and the 13BM-C station of GeoSoilEnviroCARS. High-purity  $\alpha$ -FeOOH powders were purchased from Alfa Aesar (MFCD00064782, 99+% purity). The powder sample was pre-compressed to 35  $\mu$ m (length)  $\times$  35  $\mu$ m (width)  $\times$  10  $\mu$ m (thickness) and quickly loaded into the DAC without annealing. Diamond anvils with culet diameter of 150  $\mu$ m (bevelled from 300  $\mu$ m) were used to reach high pressures. The sample was placed in a hole of diameter 95  $\mu$ m, and sealed with rhenium gasket. Neon was loaded as hydrostatic pressure medium and thermal insulator. The pressure after gas-loading was 0.9 GPa, derived from the lineshift of ruby fluorescence.

During the compression process, pressure was first calibrated by ruby fluorescence and cross-checked by diffraction peaks of solid neon<sup>34</sup> and the rhenium gasket edge. At higher pressures, while ruby fluorescence modes were hard to detect, the lowest derivative value of diamond Raman peak was taken as additional calibration. After laser heating, diffraction peaks of neon were shifted, owing to the presence of hydrogen, and could not be used as a reliable pressure calibrate. Therefore only the lowest derivative of the diamond Raman peak and diffraction pattern taken at the edge of rhenium gasket were considered for pressure calibration. The pressure uncertainty is as large as  $\pm 5$  GPa.

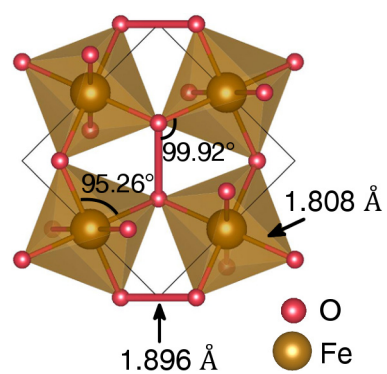
XRD patterns were collected at stations 13BM-C of GeoSoilEnviroCARS and 16-IDB of HPCAT. FeOOH powder on Ne was compressed to 92 GPa (compression rate  $\sim 10$  GPa per hour). The compressed sample was laser heated for 10 min. Sample pressure decreased to 87 GPa after quenching to ambient temperature. A Raman peak around 5,180  $\text{cm}^{-1}$  (Fig. 3) associated with the Q<sub>1</sub> vibron of hydrogen was observed, showing the diffusion of hydrogen into neon crystals<sup>15,35</sup>. A portion of the heated sample had transformed into FeO<sub>2</sub> (Fig. 2b and Extended Data Fig. 4).

Additional *in situ* laser heating experiments were conducted to investigate the stability field of FeOOH. From 58–114 GPa, the sample was heated by a double-sided laser and we probe the P-phase by both XRD and Raman spectroscopy. We constrained the decomposition pressure of FeOOH between 78 GPa and 87 GPa (Extended Data Fig. 3b).

**Ab initio crystal structure searching.** The first-principles calculations were performed in the framework of density functional theory<sup>36,37</sup> through package VASP<sup>38</sup>. The generalized gradient approximation of Perdew, Burke, and Ernzerhof was implemented to describe the exchange correlation functions<sup>39,40</sup>. Pseudopotentials were used with eight valence electrons for Fe atoms ( $3d^7 4s^1$ ) and six for O atoms ( $2s^2 2p^4$ ). Energy convergence and  $k$ -points mesh. For the crystal structure search, we used USPEX<sup>20</sup> with a plane-wave basis set cutoff energy of 800 eV.

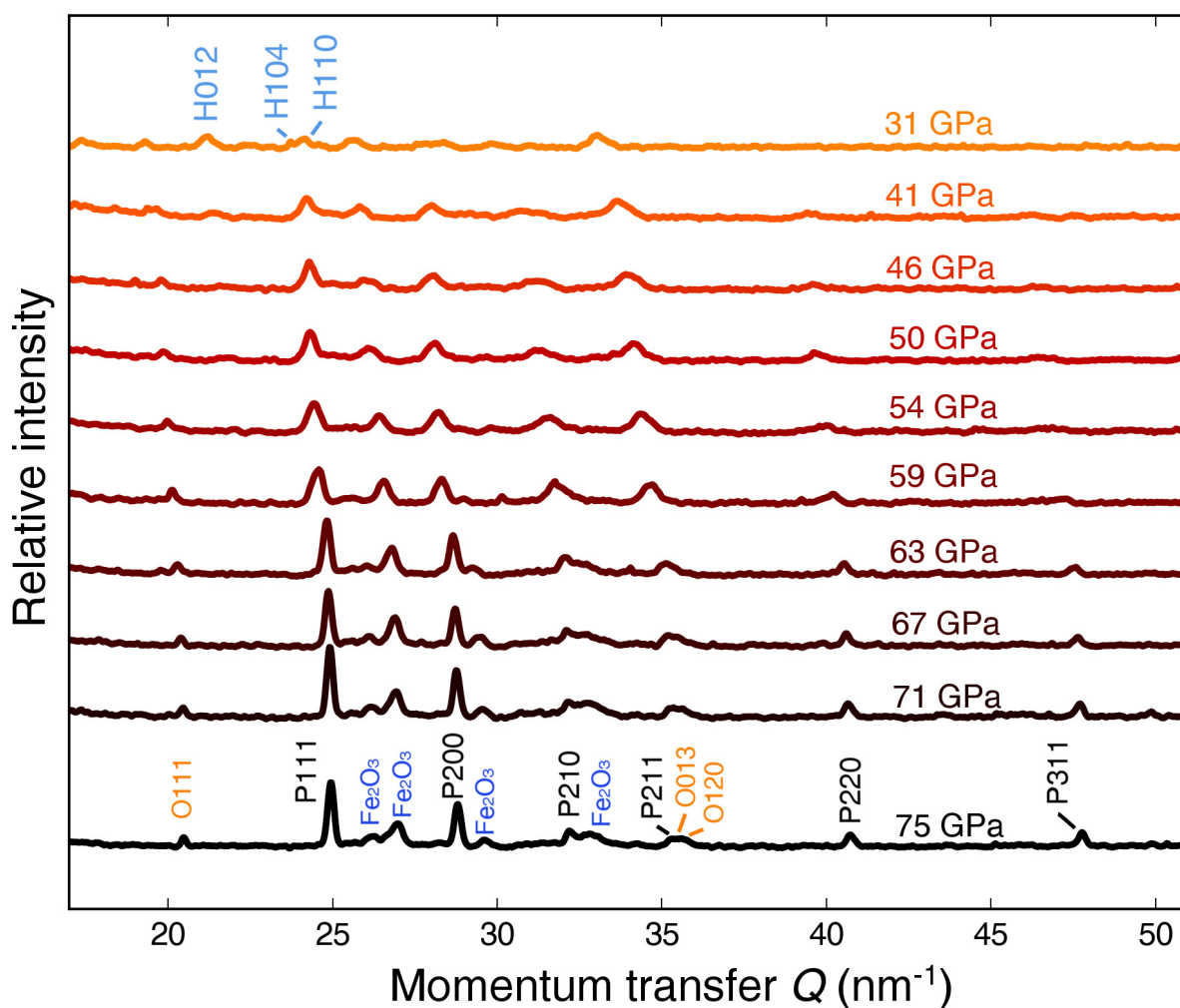
**Phonon dispersion curves from first-principles.** Phonon calculations were conducted based on density functional perturbation theory<sup>41</sup> implemented in VASP software in connection with the Phonopy software<sup>42</sup>.  $8 \times 8 \times 8$   $q$ -mesh (phonon momentum space) was used for mapping a  $2 \times 2 \times 2$  supercell of FeO<sub>2</sub>. We found that phonon dispersions are also stable at 0 GPa and 300 GPa, as shown in Extended Data Fig. 5.

33. Shim, S.-H. & Duffy, T. S. Raman spectroscopy of Fe<sub>2</sub>O<sub>3</sub> to 62 GPa. *Am. Mineral.* **87**, 318–326 (2002).
34. Fei, Y. et al. Toward an internally consistent pressure scale. *Proc. Natl Acad. Sci. USA* **104**, 9182–9186 (2007).
35. Loubeyre, P., Letoullec, R. & Pinceaux, J. P. Raman measurements of the vibrational properties of H<sub>2</sub> as a guest molecule in dense helium, neon, argon, and deuterium systems up to 40 GPa. *Phys. Rev. B* **45**, 12844–12853 (1992).
36. Hohenberg, P. & Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **136**, B864–B871 (1964).
37. Kohn, W. & Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133–A1138 (1965).
38. Perdew, J. P. et al. Atoms, molecules, solids, and surfaces: applications of the generalized gradient approximation for exchange and correlation. *Phys. Rev. B* **46**, 6671–6687 (1992).
39. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
40. Baroni, S., de Gironcoli, S., Dal Corso, A. & Giannozzi, P. Phonons and related crystal properties from density-functional perturbation theory. *Rev. Mod. Phys.* **73**, 515–562 (2001).
41. Togo, A. & Tanaka, I. First principles phonon calculations in materials science. *Scr. Mater.* **108**, 1–5 (2015).

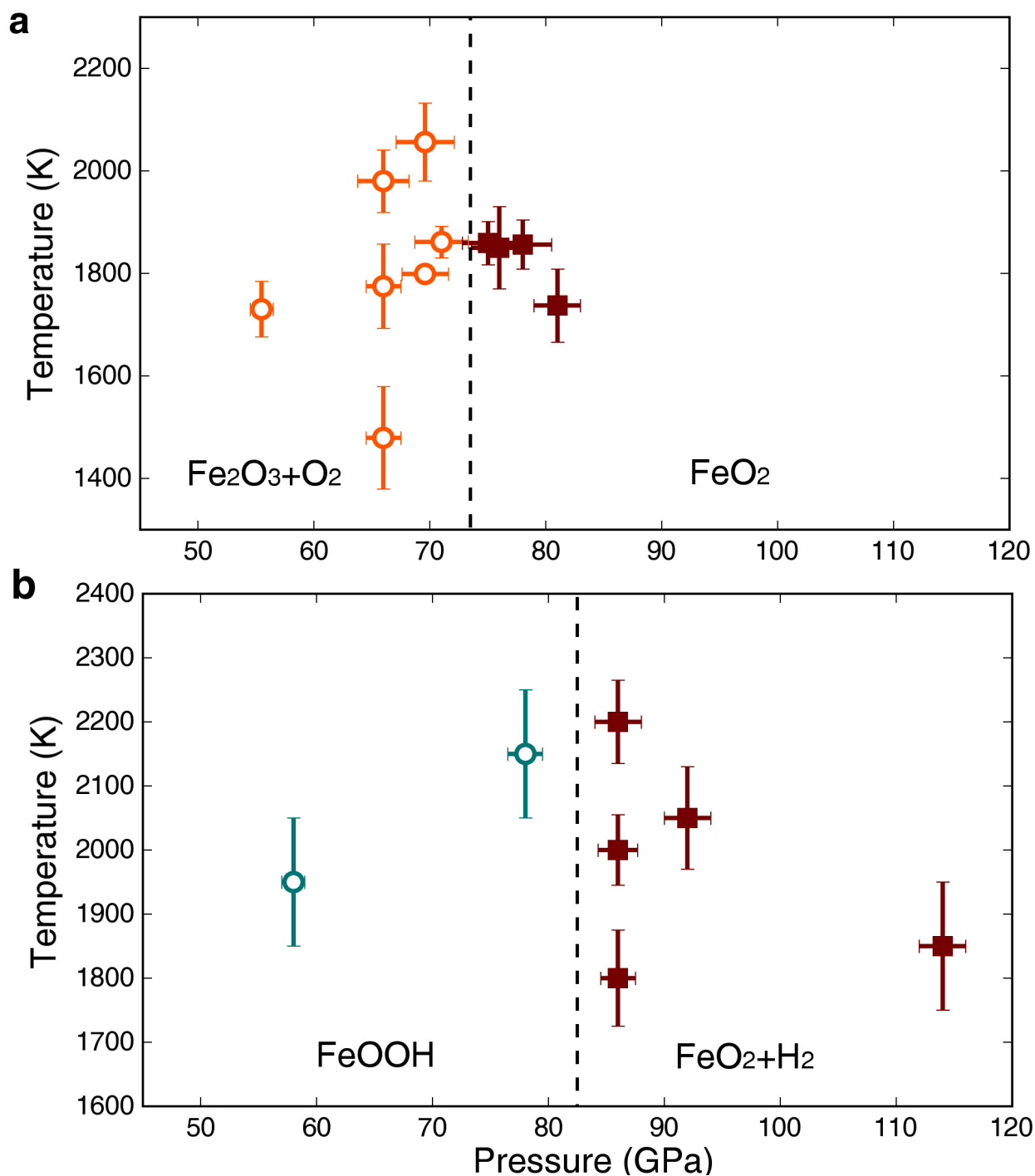


**Extended Data Figure 1 | Bonding lengths and angles in pyrite-type  $\text{FeO}_2$  at 76 GPa.** The structure is viewed along the  $x$  axis of the experimental  $\text{FeO}_2$  unit cell.



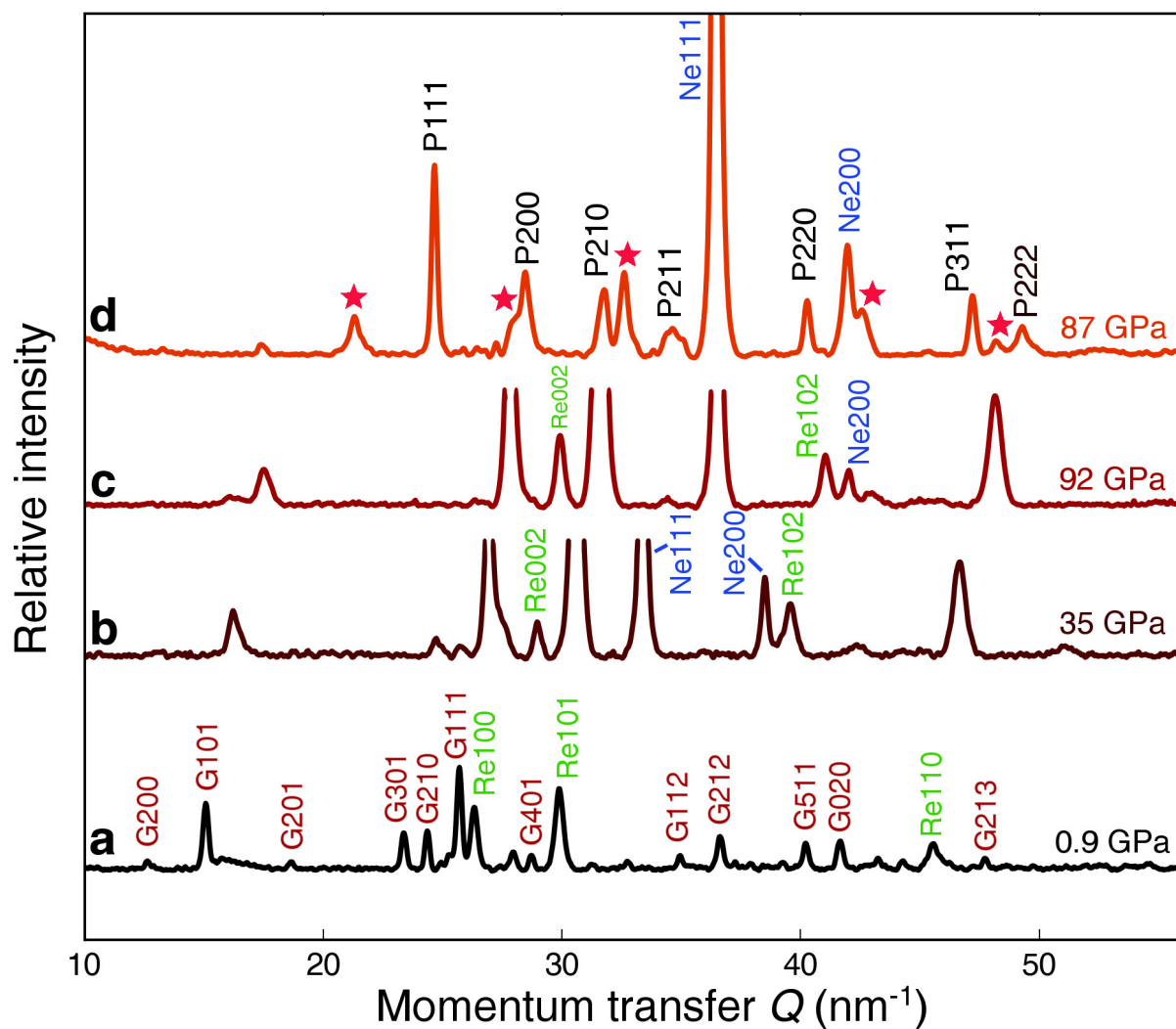


**Extended Data Figure 2 | XRD pattern series by decompressing the P-phase.** The P-phase becomes weak in intensity at 41 GPa and totally disappears at 31 GPa and below. The decompressed sample eventually recovered to the  $\alpha$ - $\text{Fe}_2\text{O}_3$  phase at low pressure. P indicates the P-phase  $\text{Fe}_2\text{O}_3$ ; O indicates solid  $\text{O}_2$ ; H indicates  $\alpha$ - $\text{Fe}_2\text{O}_3$  (haematite).  $\text{Fe}_2\text{O}_3$  contains post-perovskite type and *Aba2*-structured high-pressure phases.



**Extended Data Figure 3 | Synthesis pressure–temperature conditions for FeO<sub>2</sub>.** **a**, Open circles indicate the coexistence of Fe<sub>2</sub>O<sub>3</sub> and O<sub>2</sub>. Solid squares indicate the appearance of P-phase FeO<sub>2</sub>. FeO<sub>2</sub> was synthesized between 72 GPa and 75 GPa. **b**, Open circles indicate FeOOH.

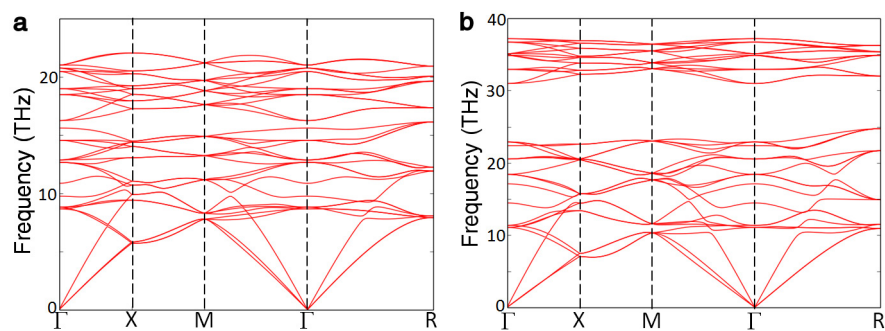
Decomposition pressure was constrained between 78 GPa and 87 GPa. Sample temperature was measured using the spectroradiometric method and errors are estimated from the goodness of fit to the spectroradiometric profile.



**Extended Data Figure 4 | XRD patterns of FeOOH through the experimental pressure–temperature path.** **a**, Goethite sample (G) in neon and Re gasket at 0.9 GPa. **b**, As in **a**, compressed to 35 GPa. The XRD peaks include goethite, solidified Ne, and the Re gasket. **c**, As in **a**, compressed to 92 GPa. The sample peaks remain and shift to higher  $Q$ .

**d**, After laser-heating and quenching to ambient temperature, the pressure dropped to 87 GPa, and the goethite peaks disappeared. The new pattern consists of peaks of the P-phase, Ne and minor amounts of  $\epsilon$ -FeOOH (red stars).





**Extended Data Figure 5 | Phonon dispersion relations of  $\text{FeO}_2$  P-phase. a, At 0 GPa; b, At 300 GPa. The P-phase is mechanically stable at 0 GPa and 300 GPa.**

Extended Data Table 1 | Additional multigrain XRD data was obtained at 13BM-C

**a**

<b>h</b>	<b>k</b>	<b>l</b>	<b>2<math>\theta</math> (°)</b>	<b><math>\omega</math> (°)</b>	<b><math>\eta</math> (°)</b>	<b><math>d_{obs}</math> (Å)</b>	<b><math>d_{cal}</math> (Å)</b>	<b><math>\Delta d/d</math></b>
0	-2	0	11.43	-11.1	30.5	2.181	2.181	0.000
0	2	0	11.44	11.3	210.4	2.181	2.181	0.000
0	-2	1	12.79	6.9	5.9	1.952	1.951	0.001
-1	-1	-3	19.03	-7.5	103.5	1.314	1.315	0.000
1	1	3	19.02	12.0	283.5	1.315	1.315	0.000

**b**

<b>h</b>	<b>k</b>	<b>l</b>	<b>2<math>\theta</math> (°)</b>	<b><math>\omega</math> (°)</b>	<b><math>\eta</math> (°)</b>	<b><math>d_{obs}</math> (Å)</b>	<b><math>d_{cal}</math> (Å)</b>	<b><math>\Delta d/d</math></b>
-1	1	1	9.88	4.9	281.5	2.522	2.520	0.001
1	-1	-1	9.89	-5.2	101.3	2.520	2.520	0.000
0	-2	0	11.42	-5.9	46.3	2.183	2.183	0.000
0	2	0	11.42	9.8	226.3	2.183	2.183	0.000
-2	0	2	16.18	10.0	317.2	1.541	1.543	-0.001

**c**

<b>h</b>	<b>k</b>	<b>l</b>	<b>2<math>\theta</math> (°)</b>	<b><math>\omega</math></b>	<b><math>\eta</math> (°)</b>	<b><math>d_{obs}</math> (Å)</b>	<b><math>d_{cal}</math> (Å)</b>	<b><math>\Delta d/d</math></b>
-1	-1	1	9.89	-8.6	82.5	2.520	2.520	0.000
1	1	-1	9.88	1.4	262.7	2.521	2.520	0.001
-1	-2	0	12.78	-0.5	120.8	1.952	1.952	0.000
2	2	-2	19.85	6.4	262.5	1.260	1.260	0.000

**d**

<b>h</b>	<b>k</b>	<b>l</b>	<b>2<math>\theta</math> (°)</b>	<b><math>\omega</math> (°)</b>	<b><math>\eta</math> (°)</b>	<b><math>d_{obs}</math> (Å)</b>	<b><math>d_{cal}</math> (Å)</b>	<b><math>\Delta d/d</math></b>
-1	1	-1	9.89	0.0	225.1	2.520	2.519	0.000
0	-2	1	12.78	-4.9	83.7	1.952	1.951	0.001
0	2	-1	12.79	8.0	263.7	1.950	1.951	0.000
-1	-3	1	19.01	-5.2	103.0	1.315	1.315	0.000

**e**

<b>h</b>	<b>k</b>	<b>l</b>	<b>2<math>\theta</math> (°)</b>	<b><math>\omega</math></b>	<b><math>\eta</math> (°)</b>	<b><math>d_{obs}</math> (Å)</b>	<b><math>d_{cal}</math> (Å)</b>	<b><math>\Delta d/d</math></b>
1	-1	-1	9.9	3.4	143.0	2.517	2.518	0.000
-1	-2	0	12.79	-7.5	68.5	1.950	1.950	0.000
1	2	0	12.8	6.5	248.6	1.949	1.950	-0.001
-1	3	1	19	6.0	296.5	1.316	1.315	0.001

**a**, Single-crystal XRD peaks for a selected P-phase crystalline with unit-cell parameter  $a = 4.3618(6)$  Å at 76 GPa (space group  $P\bar{6}3$ , incident beam wavelength  $\lambda = 0.4344$  Å). The Bragg angle  $2\theta$ , rotation angle  $\omega$  and azimuthal angle  $\eta$  are calculated from the orientation matrix.

**b–e**, XRD indexing for the second, third, fourth and fifth P-phase crystalline, with unit-cell parameters  $a = 4.3649(8)$  Å,  $4.3640(8)$  Å,  $a = 4.3625(7)$  Å,  $a = 4.3616(8)$  Å, respectively. The  $d$ -spacings are as follows:  $d_{obs}$  are the observed peak positions,  $d_{cal}$  are calculated from lattice parameters,  $\Delta d = d_{obs} - d_{cal}$ . The estimated standard deviations are included in parentheses.

Extended Data Table 2 | The lattice parameters and atomic positions of pyrite-type FeO<sub>2</sub> at 76 GPa

	$a$ (Å)	Volume (Å <sup>3</sup> )
Grain1	4.3618(6)	82.98(4)
Grain2	4.3649(8)	83.16(6)
Grain3	4.3640(8)	83.11(5)
Grain4	4.3625(7)	83.03(4)
Grain5	4.3616(8)	82.97(6)
Powder	4.3628(1)	83.04(1)
Simulation	4.331	81.23

Lattice parameters are calculated from five selected grains with X-ray wavelength 0.4344 Å. Diffraction pattern is treated and averaged as powder rings for Rietveld refinements. The Fe atoms are in 4a Wyckoff positions with fractional coordinates (0.0, 0.0, 0.0) (1/2, 0, 1/2) (0, 1/2, 1/2) (1/2, 1/2, 0) and oxygen in 8c positions ( $u, u, u$ ), ( $-u + 1/2, -u, u + 1/2$ ), ( $-u, u + 1/2, -u + 1/2$ ), ( $u + 1/2, -u + 1/2, -u$ ), ( $-u, -u, -u$ ), ( $u + 1/2, u, -u + 1/2$ ), ( $u, -u + 1/2, u + 1/2$ ), ( $-u + 1/2, u + 1/2, u$ ) with  $u = 0.3746(1)$ . Simulation lattice parameters are extracted from the lowest-energy structure search by USPEX.



Extended Data Table 3 | Bonding lengths and angles in pyrite-type FeO<sub>2</sub> at 76 GPa

	Experiment	Simulation
Pressure (GPa)	76	76
Temperature (K)	297	0
<Fe-O> (Å)	1.792(5)	1.781
<O-O> (Å)	1.937(11)	2.077
Bond angle: O-Fe-O (°)	95.6(1)	96.52
Bond angle: Fe-O-O (°)	99.1(3)	96.82

Both experimental and computational bonding information is summarized. Uncertainties are calculated from structural refinement.

# *Homo floresiensis*-like fossils from the early Middle Pleistocene of Flores

Gerrit D. van den Bergh<sup>1\*</sup>, Yousuke Kaifu<sup>2\*</sup>, Iwan Kurniawan<sup>3</sup>, Reiko T. Kono<sup>2</sup>, Adam Brumm<sup>4,5</sup>, Erick Setiyabudi<sup>3</sup>, Fachroel Aziz<sup>3</sup> & Michael J. Morwood<sup>1‡</sup>

The evolutionary origin of *Homo floresiensis*, a diminutive hominin species previously known only by skeletal remains from Liang Bua in western Flores, Indonesia, has been intensively debated. It is a matter of controversy whether this primitive form, dated to the Late Pleistocene, evolved from early Asian *Homo erectus* and represents a unique and striking case of evolutionary reversal in hominin body and brain size within an insular environment<sup>1–4</sup>. The alternative hypothesis is that *H. floresiensis* derived from an older, smaller-brained member of our genus, such as *Homo habilis*, or perhaps even late *Australopithecus*, signalling a hitherto undocumented dispersal of hominins from Africa into eastern Asia by two million years ago (2 Ma)<sup>5,6</sup>. Here we describe hominin fossils excavated in 2014 from an early Middle Pleistocene site (Mata Menge) in the So'a Basin of central Flores. These specimens comprise a mandible fragment and six isolated teeth belonging to at least three small-jawed and small-toothed individuals. Dating to ~0.7 Ma, these fossils now constitute the oldest hominin remains from Flores<sup>7</sup>. The Mata Menge mandible and teeth are similar in dimensions and morphological characteristics to those of *H. floresiensis* from Liang Bua. The exception is the mandibular first molar, which retains a more primitive condition. Notably, the Mata Menge mandible and molar are even smaller in size than those of the two existing *H. floresiensis* individuals from Liang Bua. The Mata Menge fossils are derived compared with *Australopithecus* and *H. habilis*, and so tend to support the view that *H. floresiensis* is a dwarfed descendent of early Asian *H. erectus*. Our findings suggest that hominins on Flores had acquired extremely small body size and other morphological traits specific to *H. floresiensis* at an unexpectedly early time.

This paper reports morphological analyses of hominin fossil materials excavated from the open site of Mata Menge in 2014 (ref. 7) (Extended Data Table 1). Mata Menge is one of the Middle Pleistocene fossil-bearing localities in the So'a Basin, and is situated 74 km east-southeast of Liang Bua. The specimens ( $n=7$ ) under study were recovered *in situ* from the upper part of a lens-shaped fluvial sandstone unit (Layer II) measuring up to 30 cm in thickness. Layer II is capped by a 6.5 m thick sequence of clay-rich volcanic mudflows (Layer Ia–f) that filled in the stream valley and effectively sealed off Layer II (ref. 7). All hominin fossils were excavated within a maximum linear distance of 15 m. They are associated with stone tools and the fossil remains of dwarfed proboscideans (*Stegodon florensis*), murine rodents, Komodo dragons, and other insular fauna of Flores. The age of Layer II is constrained to between 0.65 and 0.8 Ma, using <sup>40</sup>Ar/<sup>39</sup>Ar dating and other methods of age determination<sup>7</sup>. The hominin fossils display some minor dissolution pitting; generally, however, the surface preservation of these specimens is quite good.

SOA-MM4 is a right mandibular corpus (Fig. 1). Despite its small size, we conclude that this partial mandible comes from an adult individual, and that the preserved alveoli represent M<sub>1</sub> to M<sub>3</sub>. Only the lingual wall of the mesial alveolus remains for M<sub>1</sub> (Extended Data Fig. 1a). This is not for P<sub>3</sub> because the mandibular canal that normally exists in the area below P<sub>3</sub>–M<sub>1</sub> of a hominin mandible further continues anteriorly beyond this level (Extended Data Fig. 1c, h). Micro computed tomography (CT) scan data indicates that the alveolus for the last molar supported a plate-like mesial root and a conical distal root which together tilt distally, a form typical for a hominin M<sub>3</sub> root (Extended Data Fig. 1h, i). Distally to it, the alveolar bone bears no evidence of a tooth germ. The bottoms of the long M<sub>3</sub> alveoli come close to the mandibular canal and display tapering shapes, indicating that its root formation was fully or at least nearly completed.

The lateral corpus is the smallest in our sample, being 21–28% lower and narrower than in the two existing *H. floresiensis* mandibles from Liang Bua (LB1, LB6/1: Extended Data Fig. 2a, b). The lateral corporal surface of SOA-MM4 is damaged, but its cross-sectional shape (Fig. 1d, Extended Data Fig. 1e) clearly indicates the absence of an *Australopithecus*-like hollow, and the presence of a prominent superior lateral torus, a feature characteristic of *Homo*<sup>8,9</sup>. Mandibles of *Australopithecus* and to a lesser extent those of *H. habilis sensu lato* are characterized by a robust and strongly everted lateral corpus as well as a wide extramolar sulcus, in association with their narrow dental arcades and the resultant horizontal separation between the lateral mandibular corpus and the ramus<sup>10,11</sup>. These features are lacking in SOA-MM4, which has a comparatively thin, vertically oriented lateral corpus with a narrow extramolar sulcus that is evident from the medially located anterior ramus root (Extended Data Fig. 1). Such features became apparent in post-1.7-million-year-old (Myr old) *Homo*, including early Javanese *H. erectus* and *H. floresiensis* (Extended Data Fig. 3). Similarities between SOA-MM4 and the corresponding morphology of *H. floresiensis* extend to other features such as the near parallel alveolar margin and mandibular base, a moderate lateral prominence, and a gently hollowed masseteric fossa with a coarse, curved line for the masseter muscle attachment (Extended Data Fig. 4). Multivariate analyses based on the small number of the available linear measurements also support our hypothesis that SOA-MM4 is at least different from *Au. afarensis*, and is similar to *H. floresiensis* in the corpus shape (Extended Data Fig. 5).

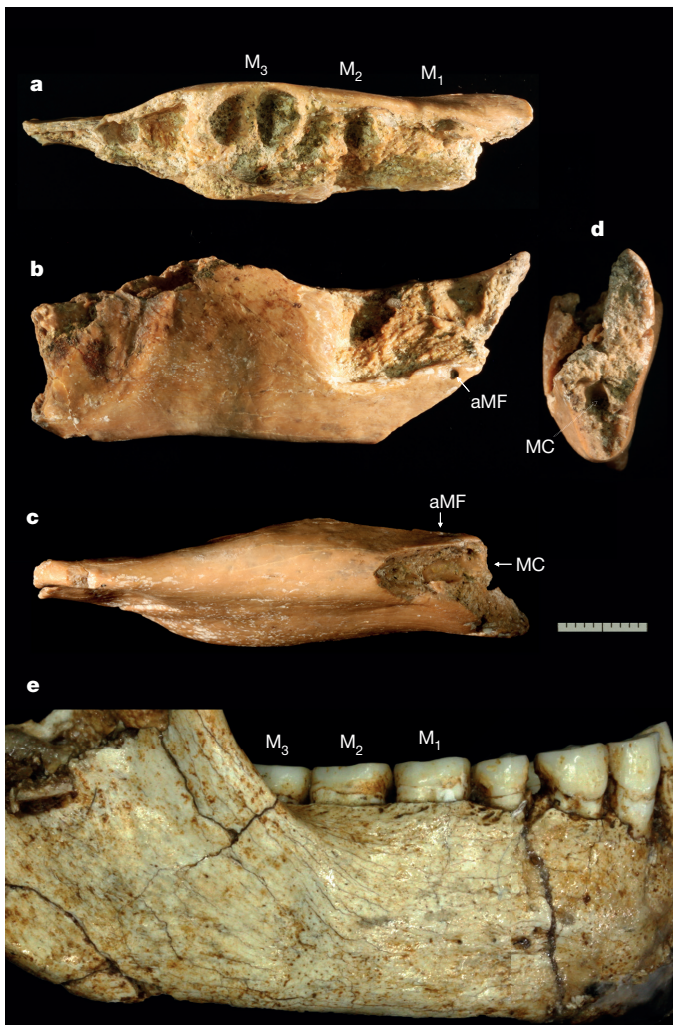
The 2014 fossil assemblage from Mata Menge includes six isolated hominin teeth from three or more individuals (Fig. 2; Extended Data Fig. 1j,k; Extended Data Table 1; Supplementary Information). Crown and root measurements available from three permanent teeth (left I<sup>1</sup>, right P<sup>3</sup>, and left M<sub>1</sub> (or M<sub>2</sub>)<sup>12</sup>) are small and similar to or slightly smaller than those of *H. floresiensis* (Extended Data Table 2,

<sup>1</sup>Centre for Archaeological Science, School of Earth & Environmental Sciences, University of Wollongong, Wollongong, New South Wales 2522, Australia. <sup>2</sup>Department of Anthropology, National Museum of Nature and Science, 4-1-1 Amakubo, Tsukuba-shi, Ibaraki 305-0005, Japan. <sup>3</sup>Geology Museum Bandung, Geological Agency, Jalan Diponegoro 57, Bandung 40122, Indonesia.

<sup>4</sup>Research Centre of Human Evolution, Environmental Futures Research Institute, Griffith University, Nathan, Queensland 4111, Australia. <sup>5</sup>School of Earth & Environmental Sciences, University of Wollongong, Wollongong, New South Wales 2522, Australia.

\*These authors contributed equally to this work.

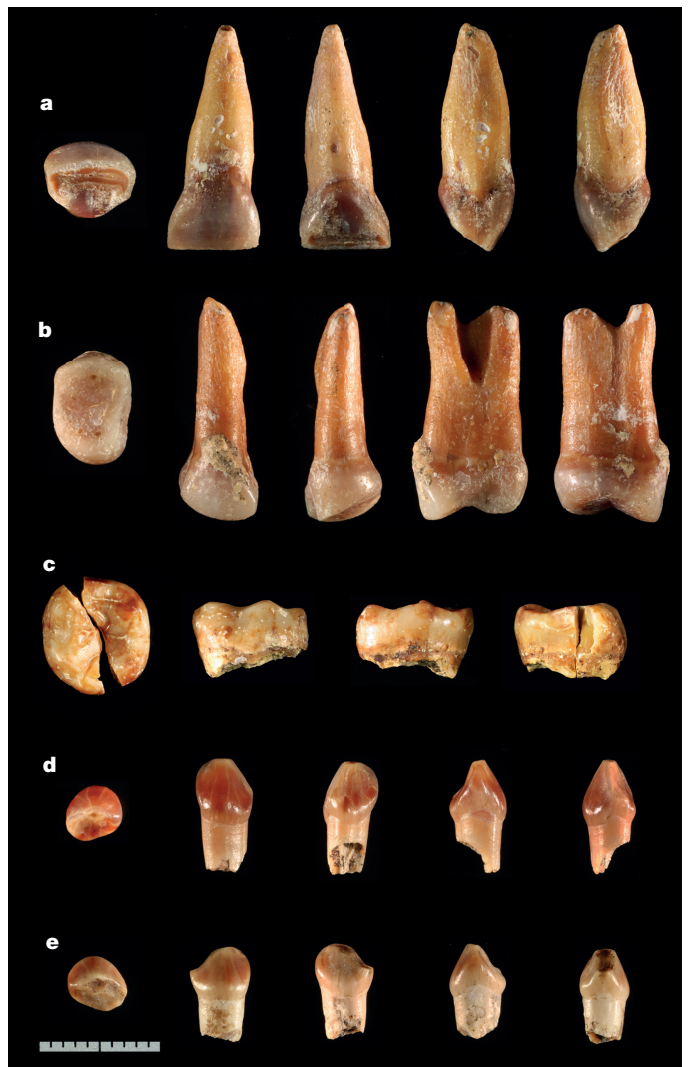
‡Deceased.



**Figure 1 | SOA-MM4 mandible compared with a Liang Bua *H. floresiensis* specimen. a–d**, Superior (a), lateral (b), inferior (c), and anterior (d) views. **e**, Lateral view of the LB6/1 mandible. M<sub>1</sub>, first molar; M<sub>2</sub>, second molar; M<sub>3</sub>, third molar; MC, mandibular canal; aMF, accessory mental foramen. Scale bar, 10 mm.

Extended Data Fig. 2c–f). The broken root of the I<sub>1/2</sub> is also equally small (Extended Data Fig. 1k), although comparative measurements are unavailable from this specimen, which was used for direct uranium-series dating<sup>7</sup>. Morphologically, the Mata Menge teeth display the following primitive features: (i) a lingually (I<sup>1</sup>, I<sub>1/2</sub>) or distally (P<sup>3</sup>) beveled, worn occlusal surface that suggests tilted anterior dentition and substantial prognathism (Extended Data Fig. 1j); (ii) a pronounced P<sup>3</sup> lingual cusp whose mesiodistal diameter compares with that of the buccal cusp<sup>13,14</sup>; and (iii) a mid-trigonid crest on M<sub>1</sub>. These features are frequently observed in Early Pleistocene African and Eurasian *Homo* (that is, *H. habilis sensu lato* and *H. erectus sensu lato*), and the third character became frequent in *H. erectus* and some later groups of archaic *Homo*<sup>3,15</sup>. Liang Bua *H. floresiensis* shares the first and probably the third characteristics, although the second is not evident on the worn Liang Bua premolars<sup>16</sup>. Most features of the Mata Menge I<sup>1</sup> and P<sup>3</sup> are not useful for assessing taxonomic affinities relative to *H. habilis* or *H. erectus* (Supplementary Information), although the absence of the P<sup>3</sup> buccal groove is a condition appeared in post-*habilis* grade *Homo*<sup>3</sup>. The Mata Menge and Liang Bua hominins also share a bifurcated, fused P<sup>3</sup> root form.

We digitally reconstructed the broken M<sub>1</sub> (or M<sub>2</sub>) crown (SOA-MM1) based on its micro-CT scan (Fig. 3a). Both linear metric and crown contour analyses of the M<sub>1</sub>s showed that this five-cusped

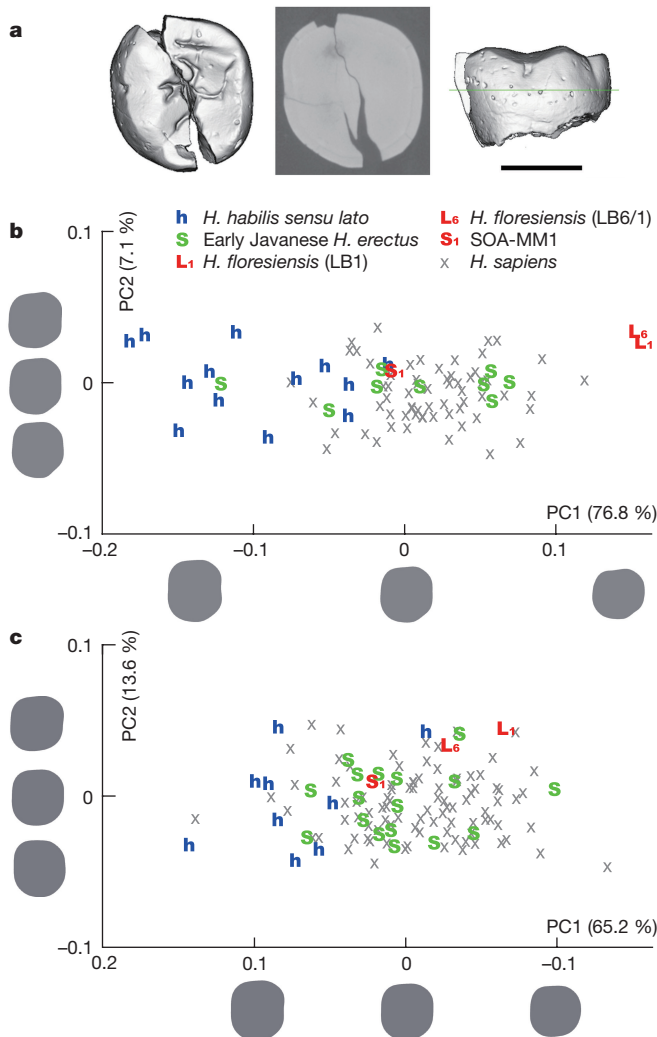


**Figure 2 | Isolated teeth from Mata Menge. a**, SOA-MM2 (left I<sup>1</sup>). **b**, SOA-MM5 (right P<sup>3</sup>). **c**, SOA-MM1 (left M<sub>1</sub>). **d**, SOA-MM7 (left d<sub>c</sub>). **e**, SOA-MM8 (right d<sub>c</sub>). In each row, from left to right, occlusal, buccal (labial), lingual, mesial, and distal (except for c) views. Scale bar, 10 mm.

tooth is moderately long and is close to the average M<sub>1</sub> shape of early Javanese *H. erectus*, but is different from the elongated *H. habilis*-like forms<sup>17</sup> (Fig. 3b, Extended Data Fig. 2e). SOA-MM1 lacks two of the most peculiar, derived characteristics of the Liang Bua *H. floresiensis* M<sub>1</sub>s: a reduced cusp number (five to four) and a MD shortened crown configuration<sup>3</sup>. The above comparative morphology remains largely the same even if SOA-MM1 is a M<sub>2</sub>, although these analyses do not clearly separate *H. floresiensis* from early Javanese *H. erectus* (Fig. 3c; Extended Data Fig. 2f).

The two deciduous canines (d<sub>c</sub>s) from Mata Menge are much smaller than *H. sapiens* ( $n = 63$ ), *H. erectus* ( $n = 1$ ), and *Australopithecus* ( $n = 6$ ), but do not display the relatively high crown shape that characterizes the latter genus (Extended Data Fig. 6a, b). In a principal component analysis (PCA) based on five size-adjusted linear measurements (Extended Data Fig. 6c–e), PC1 separates *Australopithecus*-like primitive (a high crown with a low distal shoulder) and modern human-like derived (a low crown with a high distal shoulder) morphologies. Allometry does not explain this inter-taxon difference because the d<sub>c</sub> crown sizes are similar between the two taxa. SOA-MM7, the minimally worn d<sub>c</sub> from Mata Menge, is positioned in between *Australopithecus* and *H. sapiens* in PC1. There are no deciduous teeth in the existing *H. floresiensis* assemblage from Liang Bua.





**Figure 3 | CT-based reconstruction of SOA-MM1 and the results of Elliptic Fourier Analysis of the molar crown contour.** **a**, Occlusal (left) and buccal (right) views of the reconstructed SOA-MM1, and a horizontal CT section (central) at the level indicated in the buccal view. Scale bar, 5 mm. Plots of the PC scores for the first molar (**b**) and the second molar (**c**) analyses. Proportions of the variance explained by each PC are in the parentheses.

The above findings shed new light on the origin and evolution of Late Pleistocene *H. floresiensis*. Notably, the 0.7-Myr-old Mata Menge hominins are similar to Late Pleistocene *H. floresiensis* of Liang Bua in dentognathic size and morphology, but the former lacks several derived molar morphologies of the latter. This suggests that the early Middle Pleistocene hominins of the So'a Basin were directly ancestral to Liang Bua *H. floresiensis*. Further support for this view is provided by the following observations: (i) stone technologies at Mata Menge and Liang Bua are markedly similar, implying a period of technological continuity spanning at least several hundred millennia<sup>18</sup>; (ii) there is no evidence for a faunal turnover during the time interval separating the fossil records of the So'a Basin and Liang Bua<sup>19</sup>; and (iii) *H. floresiensis* lacks a series of derived cranial features of chronologically late *H. erectus* from Java, such as specimens from Ngandong, Sambungmacan, and Ngawi (all of which are presumably from the late Middle to Late Pleistocene period)<sup>2,20</sup>. We conclude that the most reasonable taxonomic assignment for the Mata Menge fossils is to *H. floresiensis*, although this remains a provisional interpretation until new skeletal materials are found.

Concerning the origins and evolutionary relationships of *H. floresiensis*, we note that the Mata Menge mandible and teeth

are morphologically derived compared with *Australopithecus* and *H. habilis*, with their primitive aspects comparable to post-*habilis* grade Early Pleistocene *Homo*. This is most consistent with the hypothesis that *H. floresiensis* originated from a population whose closest affinities are with early Javanese *H. erectus* ( $\geq 1.2$ –0.8 Ma), whose femoral length is 55–61% longer, and absolute brain size about twice as large, as *H. floresiensis*<sup>21–23</sup>. Additional support for this includes reports that the earliest evidence for hominins on Flores ( $\sim 1.0$  Ma)<sup>24</sup> does not exceed the oldest record of *H. erectus* on Java ( $\geq 1.2$  Ma)<sup>25,26</sup>, and recent detailed analyses of the craniodental morphology of Liang Bua *H. floresiensis*<sup>2,3,16</sup>. Given how little is known about the distribution of early *H. erectus* on the ancient 'Sunda' shelf, it remains an open question whether the founding population crossed to Flores in a west-to-east direction from Java, or via a northern route from the Wallacean island of Sulawesi<sup>27–29</sup>.

It is noteworthy that the mandible and teeth from Mata Menge are slightly smaller than the two *H. floresiensis* individuals from Liang Bua (LB1 and LB6/1). While this could indicate a slight body size increase over time, it may also simply reflect intra-population variation in the Mata Menge and Liang Bua hominin groups. Whichever the case, it would appear that the Flores hominins had acquired extremely small dentognathic size during the time span of at least 300 millennia following the initial colonization of Flores, assuming that the oldest artefacts from Flores—dated to at least  $\sim 1$  Ma<sup>24</sup>—were produced by large-bodied ancestors of the Mata Menge hominins. This apparently very fast transformation in hominin body size is surprising. Although no other documented examples of rapid island dwarfing exist for primates, we note that red deer from the island of Jersey had reduced to one-sixth of the body size in the ancestral population within about six millennia<sup>30</sup>. Flores may have been an exceptional case; however, the fossil evidence from Mata Menge highlights how quickly major evolutionary changes could have occurred in hominin populations cut off on isolated and impoverished islands of Wallacea.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 30 December 2015; accepted 13 April 2016.**

1. Brown, P. *et al.* A new small-bodied hominin from the Late Pleistocene of Flores, Indonesia. *Nature* **431**, 1055–1061 (2004).
2. Kaifu, Y. *et al.* Craniofacial morphology of *Homo floresiensis*: description, taxonomic affinities, and evolutionary implication. *J. Hum. Evol.* **61**, 644–682 (2011).
3. Kaifu, Y. *et al.* Unique dental morphology of *Homo floresiensis* and its evolutionary implications. *PLoS ONE* **10**, e0141614 (2015).
4. Sutikna, T. *et al.* Revised stratigraphy and chronology for *Homo floresiensis* at Liang Bua in Indonesia. *Nature* **532**, 366–369 (2016).
5. Morwood, M. J. & Jungers, W. L. Conclusions: implications of the Liang Bua excavations for hominin evolution and biogeography. *J. Hum. Evol.* **57**, 640–648 (2009).
6. Brown, P. & Maeda, T. Liang Bua *Homo floresiensis* mandibles and mandibular teeth: a contribution to the comparative morphology of a new hominin species. *J. Hum. Evol.* **57**, 571–596 (2009).
7. Brumm, A. *et al.* Age and context of the oldest known hominin fossils from Flores. *Nature* <http://dx.doi.org/10.1038/nature17663> (2016).
8. White, T. D., Johanson, D. C. & Kimbel, W. H. *Australopithecus africanus*: its phyletic position reconsidered. *S. Afr. J. Sci.* **77**, 445–470 (1981).
9. Villmoare, B. *et al.* Early *Homo* at 2.8 Ma from Ledi-Geraru, Afar, Ethiopia. *Science* **347**, 1352–1355 (2015).
10. Rosas, A. & Bermúdez de Castro, J. M. On the taxonomic affinities of the Dmanisi mandible (Georgia). *Am. J. Phys. Anthropol.* **107**, 145–162 (1998).
11. Kaifu, Y. *et al.* Taxonomic affinities and evolutionary history of the Early Pleistocene hominids of Java: dentognathic evidence. *Am. J. Phys. Anthropol.* **128**, 709–726 (2005).
12. Suwa, G. Serial allocations of isolated mandibular molars of unknown taxonomic affinities from the Shungra and Usno Formations, Ethiopia, a combined method approach. *Hum. Evol.* **11**, 269–282 (1996).
13. Kimbel, W. H., Johanson, D. C. & Rak, Y. Systematic assessment of a maxilla of *Homo* from Hadar, Ethiopia. *Am. J. Phys. Anthropol.* **103**, 235–262 (1997).
14. Gómez-Robles, A., Martínón-Torres, M., Bermúdez de Castro, J. M., Prado-Simon, L. & Arsuaga, J. L. A geometric morphometric analysis of hominin upper premolars. Shape variation and morphological integration. *J. Hum. Evol.* **61**, 688–702 (2011).

15. Martínón-Torres, M. *et al.* Dental remains from Dmanisi (Republic of Georgia): morphological analysis and comparative study. *J. Hum. Evol.* **55**, 249–273 (2008).
16. Kaifu, Y. *et al.* Descriptions of the dental remains of *Homo floresiensis*. *Anthropol. Sci.* **123**, 129–145 (2015).
17. Tobias, P. V. *Olduvai Gorge, 4: The Skulls, Endocasts and Teeth of Homo habilis*. Vol. 4 (Cambridge Univ. Press, 1991).
18. Brumm, A. *et al.* Early stone technology on Flores and its implications for *Homo floresiensis*. *Nature* **441**, 624–628 (2006).
19. van den Bergh, G. D. *et al.* The Liang Bua faunal remains: a 95k.yr. sequence from Flores, East Indonesia. *J. Hum. Evol.* **57**, 527–537 (2009).
20. Kaifu, Y. *et al.* *Homo erectus* calvaria from Ngawi (Java) and its evolutionary implications. *Anthropol. Sci.* **123**, 161–176 (2015).
21. Puymerail, L. *et al.* Structural analysis of the Kresna 11 *Homo erectus* femoral shaft (Sangiran, Java). *J. Hum. Evol.* **63**, 741–749 (2012).
22. Ruff, C. B., Puymerail, L., Macchiarelli, R., Sipla, J. & Ciochon, R. L. Structure and composition of the Trinil femora: functional and taxonomic implications. *J. Hum. Evol.* **80**, 147–158 (2015).
23. Kubo, D., Kono, R. T. & Kaifu, Y. Brain size of *Homo floresiensis* and its evolutionary implications. *Proc. R. Soc. Lond. B* **280**, 20130338 (2013).
24. Brumm, A. *et al.* Hominins on Flores, Indonesia, by one million years ago. *Nature* **464**, 748–752 (2010).
25. Larick, R. *et al.* Early Pleistocene  $^{40}\text{Ar}/^{39}\text{Ar}$  ages for Bapang Formation hominins, Central Java, Indonesia. *Proc. Natl Acad. Sci. USA* **98**, 4866–4871 (2001).
26. Hyodo, M. *et al.* High-resolution record of the Matuyama–Brunhes transition constrains the age of Javanese *Homo erectus* in the Sangiran dome, Indonesia. *Proc. Natl Acad. Sci. USA* **108**, 19563–19568 (2011).
27. Morwood, M. & van Oosterzee, P. *The Discovery of the Hobbit* (Random House, 2007).
28. van den Bergh, G. D. *et al.* Earliest hominin occupation of Sulawesi, Indonesia. *Nature* **529**, 208–211 (2016).
29. Dennell, R. W., Louys, J., O'Regan, H. J. & Wilkinson, D. M. The origins and persistence of *Homo floresiensis* on Flores: biogeographical and ecological perspectives. *Quat. Sci. Rev.* **96**, 98–107 (2014).
30. Lister, A. M. Rapid dwarfing of red deer on Jersey in the last interglacial. *Nature* **342**, 539–542 (1989).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** Funding for the So'a Basin project was provided by an Australian Research Council (ARC) *Discovery* grant (DP1093342) awarded to M.J.M. and A.B., and the project was directed by M.J.M. (2010–2013) and G.D.v.d.B. (2013–2015). The Japan Society for the Promotion of Science provided a grant (No. 24247044) to Y.K. Financial and technical support was provided by the Geological Survey Centre of Indonesia. The Indonesian State Ministry of Research and Technology granted permission to undertake this research, and we thank the successive directors of the Geological Survey Centre, Y. Kusumahbrata, A. Wibowo and A. Pribadi, the Heads of the Geological Agency (R. Sukyar and Surono), and the successive directors of the Geology Museum in Bandung (S. Baskoro and O. Abdurahman) for facilitating and supporting this research. In addition, we acknowledge support and advice provided by I. Setiadi, D. Pribadi and Suyono. We also thank M. R. Puspaningrum, H. Insani, I. Sutisna, S. Sonjaya, U. P. Wibowo, A. Gunawan, A. M. Saiful, S. Hayes, B. Burhan, E. Sukandar, A. Rahman, A. Rahmadi and E. E. Laksmana for their assistance in the field (2014–2015), and G. Suwa, T. Djubiantono, F. Aziz, T. Jacob, E. Mbua, F. Schrenk, I. Tattersall, K. Mowbray, J. de Vos, P. Mennecier, F. Demeter, Nguyen Kim Thuy, and Nguyen Lan Cuong for access to the specimens in their care.

**Author Contributions** G.D.v.d.B. is co-director of the So'a Basin project with A.B. and I.K. M.J.M. (deceased) was also a co-director of the So'a Basin project. The project was initiated by M.J.M. and F.A. G.D.v.d.B. and I.K. undertook initial identification and analyses of the hominin fossils. Y.K. conducted the morphological analysis together with R.T.K. Fieldwork was planned and directed by G.D.v.d.B., I.K., A.B. and E.S.. Y.K., G.D.v.d.B. and A.B. wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Y.K. ([kaifu@kahaku.go.jp](mailto:kaifu@kahaku.go.jp)).

## METHODS

**Comparative samples.** Comparative fossil samples include the proposed two major ancestral candidates for *H. floresiensis*, *H. habilis sensu lato* (East Africa) and early Javanese *H. erectus* (Java), as well as other Asian archaic *Homo* (Dmanisi *Homo* and Middle Pleistocene Asian *Homo*), and *H. floresiensis* from the Late Pleistocene of Liang Bua (Extended Data Table 3). No *Australopithecus* molar samples were included in the comparison with SOA-MM1 because the latter is obviously derived in having a well-developed mid-trigonid crest<sup>15</sup> and a gently convex, non-bilobed, buccal crown outline<sup>8</sup>. However, in view of the previous claim that the *H. floresiensis* mandibles resemble *Australopithecus afarensis* (ref. 6, but see ref. 2), the mandibular analysis includes specimens belonging to this species as well as the recently reported 'earliest *Homo*' specimen from Ethiopia<sup>9</sup>. The deciduous tooth analyses also include *Australopithecus* specimens because no measurable mandibular deciduous canines are represented in the existing fossil collections of *H. habilis* to represent the primitive condition in this tooth. Two worn *H. erectus* specimens from the Zhoukoudian Lower Cave are included in the metrical comparison, but could not be included in the PCA; nor are there any Javanese *H. erectus* specimens of this tooth known.

Debate continues over whether *H. habilis sensu lato* includes diverse evolving lineages<sup>31–34</sup>, but we pooled the relevant specimens from East Africa for the present purpose to recognize primitive morphological condition in *Homo*.

The early Javanese *H. erectus* sample is from the varying stratigraphic levels at Sangiran, Central Java, which are dated to between  $\geq 1.2$  and 0.8 Ma<sup>25,26</sup>. Previous studies demonstrated significant temporal decrease in tooth crown size within this sample<sup>11,35,36</sup>, but we nevertheless pooled these chronological subgroups for the purposes of this study because their crown shapes are remarkably similar to each other<sup>3</sup>.

The other Asian archaic *Homo* samples (such as Dmanisi *Homo*, and various groups of the Middle Pleistocene East Asian *Homo*, as listed in Extended Data Table 3) were included in the linear metric analyses of the mandible and teeth.

Our *H. sapiens* sample is from Africa, Europe, Asia, and Oceania, with particular emphasis on prehistoric individuals from Southeast Asia, including Flores, as well as modern small-bodied populations (such as Philippine 'Negrito', Andaman, African 'Pygmy', and 'Bushman') (Extended Data Table 4). This choice was made to reflect species-wide variation of *H. sapiens*, and to respond to the claim that Liang Bua *H. floresiensis* resembles a local short-statured Australomelanesian population<sup>37</sup>. Sexes were pooled due to general difficulties in sex assignment for various fragmentary hominin fossils.

**Materials.** The data of the mandibles were taken from the original specimens, but some notes should be made on the materials for the dental analyses. Dental specimens with severe tooth wear were excluded from the metric analyses. Metric and non-metric data were obtained from the original specimens, plaster casts, or published studies (Extended Data Table 3). For all the *H. floresiensis*, early Javanese *H. erectus*, and *H. sapiens* specimens, high-quality 'isolated' plaster casts were prepared by Y.K. with partial assistance from Hisao Baba. Silicone was used for molding and the produced plaster cast of a dentition was then cut with a saw to isolate individual teeth. Such isolated casts can be measured more easily and accurately than the original specimens when the teeth are embedded in the jaw bones and measurement equipment is difficult to apply to the original specimens. Thus, we used these isolated casts for linear measurement and crown contour extraction. Non-isolated, high-quality plaster casts were used for most of the *H. habilis* and *H. ergaster* specimens. These were prepared by Gen Suwa with dimensional accuracies being within  $\pm 0.1$  mm<sup>38</sup>.

**Measurements.** A digital sliding caliper (Mitutoyo) was used for linear measurements. Mesiodistal (MD) and buccolingual (BL) tooth crown diameters were recorded with allowance for wear, following the methods outlined in ref. 39. Values from the right and left sides are averaged for the fossil specimens, while the data for *H. sapiens* are from the better-preserved and/or less-worn side. All the metric data were taken by Y.K., with the exception of those cited from the literature (refs. 15, 39–42).

**CT scan.** The Mate Menge hominin fossils reported here were CT scanned using the microfocal X-ray CT system TXS320-ACTIS (TESCO, Japan) at the National Museum of Nature and Science, Tokyo, in 2014. Original scans were taken at 189 kV and 0.23 mA with a 0.5-mm-thick copper plate prefilter to lessen beam-hardening effects. Scanned images were reconstructed into a 512 × 512 matrix of 150 µm pixel size with 150 µm slice interval and thickness (mandible), or 512 × 512 matrices of 32 µm pixels with 32 microns slice interval and 34.63 microns slice thickness (isolated teeth).

**Size-adjusted PCAs of the mandibular corpus measurements.** Principal component analysis (PCA) was performed using mandibular corpus heights and widths as variables (Extended Data Fig. 5). The size adjustment was done by dividing each raw measurement by the geometric mean of all the measurements used for each individual.

**Elliptic Fourier analysis (EFA) of the mandibular molar.** Occlusal crown contours of the mandibular molar were analysed by normalized (size-standardized) EFA<sup>43,44</sup> (Fig. 3), following the previous analysis of the teeth of *H. floresiensis*<sup>3</sup>. This method was chosen in that study because the *H. floresiensis* teeth are moderately

worn and retain few homologous landmarks. We performed EFAs for both M<sub>1</sub> and M<sub>2</sub>s, because the position of SOA-MM1 is indeterminate.

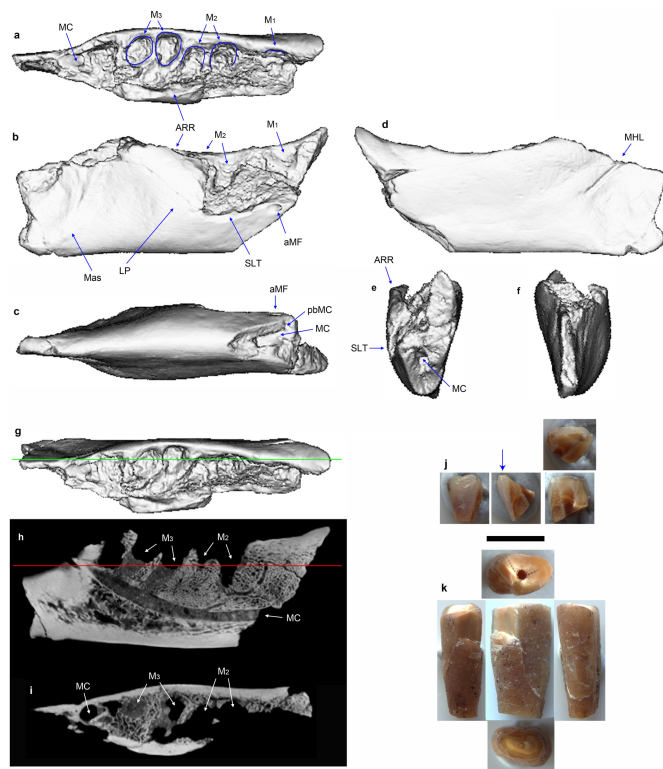
The comparative samples were *H. floresiensis*, the two major ancestral candidates of *H. floresiensis* (*H. habilis sensu lato* and early Javanese *H. erectus*), and *H. sapiens*. Comparisons were made on the images from the right side teeth, or horizontally flipped images of the left teeth if the latter side is better preserved. The crown contour of each tooth was captured by photography with a dental cast placed so that its cervical line is vertical to the axis of the camera lens<sup>45–47</sup>. Local fluctuations of the cervical lines were ignored<sup>38</sup>. A 100 mm macro lens was set to a Canon D40 digital camera to minimize the parallax effect. Interproximal wear was corrected on each photograph before extracting the crown contour. Capturing of crown contours from the digital images, obtaining EFDs, and PCA of the normalized EFDs were conducted using the software SHAPE 1.3 (ref. 48). Other methodological details for the EFA are available in ref. 3.

**Size-adjusted PCAs of the mandibular deciduous canine (d<sub>c</sub>).** Principal component analysis (PCA) was undertaken using five size-adjusted linear crown diameters (mesiodistal diameter, buccolingual diameter, mesial crown shoulder height, crown height, and distal crown shoulder height), as shown in Extended Data Fig. 6c–e. The size adjustment was done by dividing each raw measurement by the geometric mean of the five measurements for each individual. Crown height of the less worn SOA-MM7 can be estimated with some confidence, but moderately worn SOA-MM8 was excluded from this analysis.

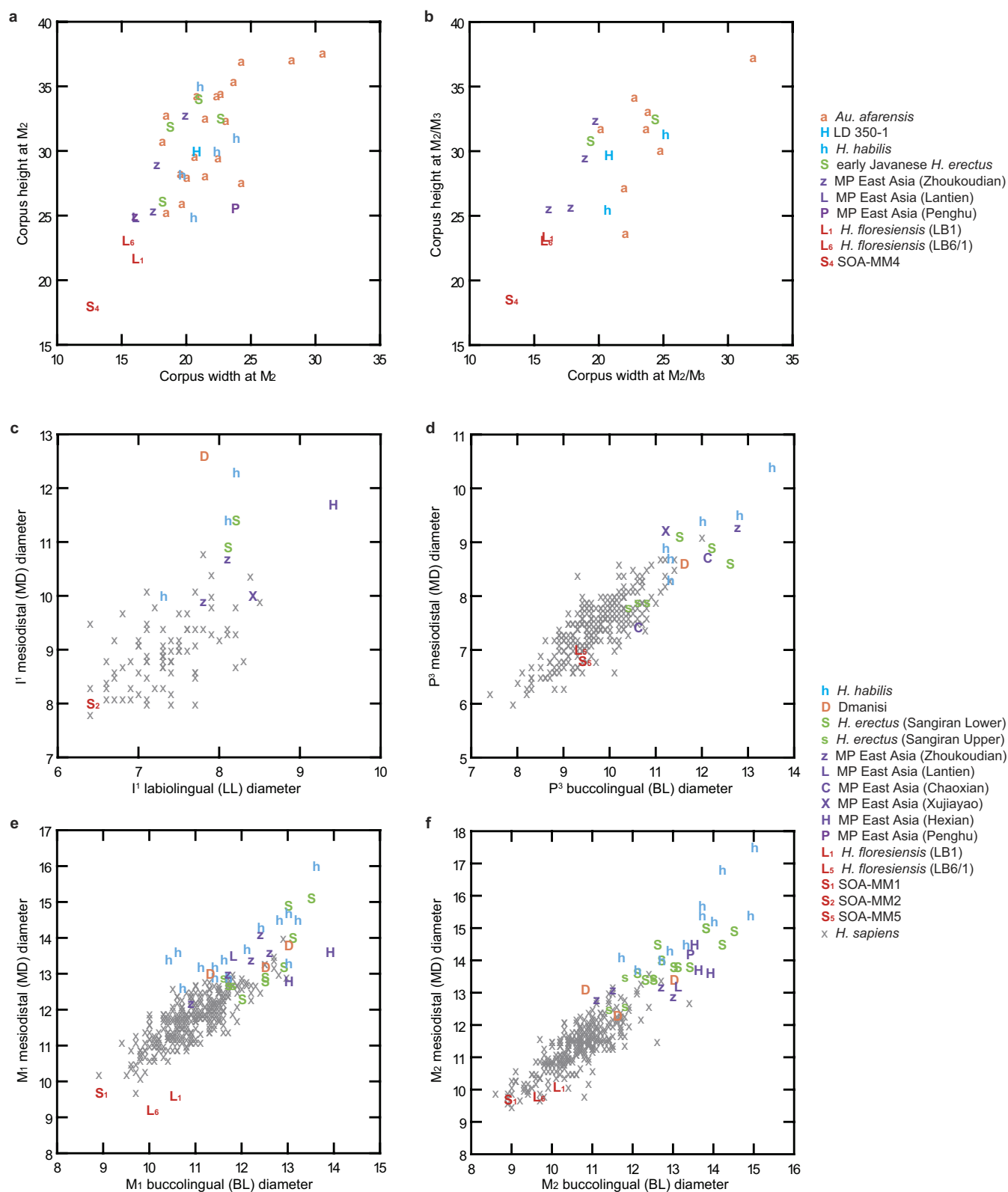
**Data reporting.** No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

- Lordkipanidze, D. et al. A complete skull from Dmanisi, Georgia, and the evolutionary biology of early *Homo*. *Science* **342**, 326–331 (2013).
- Wood, B. Fifty years after *Homo habilis*. *Nature* **508**, 31–33 (2014).
- Antón, S. C., Potts, R. & Aiello, L. C. Human evolution. Evolution of early *Homo*: an integrated biological perspective. *Science* **345**, 1236828 (2014).
- Spoor, F. et al. Reconstructed *Homo habilis* type OH 7 suggests deep-rooted species diversity in early *Homo*. *Nature* **519**, 83–86 (2015).
- Kaifu, Y. Advanced dental reduction in Javanese *Homo erectus*. *Anthropol. Sci.* **114**, 35–43 (2006).
- Zanolli, C. Additional evidence for morpho-dimensional tooth crown variation in a new Indonesian *H. erectus* sample from the Sangiran Dome (Central Java). *PLoS ONE* **8**, e67233 (2013).
- Jacob, T. et al. Pygmoid Australomelanesian *Homo sapiens* skeletal remains from Liang Bua, Flores: population affinities and pathological abnormalities. *Proc. Natl Acad. Sci. USA* **103**, 13421–13426 (2006).
- Suwa, G. A Comparative Analysis of Hominid Dental Remains from the Sungura and Usno Formations, Omo Valley, Ethiopia Ph.D. thesis, University of California, Berkeley (1990).
- Wood, B. *Koobi Fora Research Project 4: Hominid Cranial Remains*. Vol. 4 (Clarendon Press, 1991).
- Leakey, M. G. et al. New fossils from Koobi Fora in northern Kenya confirm taxonomic diversity in early *Homo*. *Nature* **488**, 201–204 (2012).
- Weidenreich, F. The dentition of *Sinanthropus pekinensis*: a comparative odontography of the hominids. *Palaeontologia Sinica New Series D* **1**, 1–180 (1937).
- Xing, S., Martínón-Torres, M., Bermúdez de Castro, J. M., Wu, X. & Liu, W. Hominin teeth from the early Late Pleistocene site of Xujiayao, Northern China. *Am. J. Phys. Anthropol.* **156**, 224–240 (2015).
- Kuhl, F. P. Elliptic Fourier features of a closed contour. *Computer Graphics and Image Processing* **18**, 236–258 (1982).
- Lestrel, P. E., Wolfe, C. A. & Bodt, A. Mandibular shape analysis in fossil hominins: Fourier descriptors in *norma lateralis*. *Homo* **64**, 247–272 (2013).
- Wood, B. A. & Abbott, S. A. Analysis of the dental morphology of Plio-Pleistocene hominids. I. Mandibular molars: crown area measurements and morphological traits. *J. Anat.* **136**, 197–219 (1983).
- Bailey, S. E. & Lynch, J. M. Diagnostic differences in mandibular P4 shape between Neanderthals and anatomically modern humans. *Am. J. Phys. Anthropol.* **126**, 268–277 (2005).
- Martínón-Torres, M. et al. Hominin lower second premolar morphology: evolutionary inferences through geometric morphometric analysis. *J. Hum. Evol.* **50**, 523–533 (2006).
- Iwata, H. & Ukai, Y. SHAPE: a computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors. *J. Hered.* **93**, 384–385 (2002).
- Smith, B. H. Patterns of molar wear in hunter-gathers and agriculturalists. *Am. J. Phys. Anthropol.* **63**, 39–56 (1984).
- Hillson, S., Fitzgerald, C. & Flinn, H. Alternative dental measurements: proposals and relationships with other measurements. *Am. J. Phys. Anthropol.* **126**, 413–426 (2005).
- Kimbel, W. H., Rak, Y. & Johanson, D. C. *The Skull of Australopithecus afarensis* (Oxford Univ. Press, 2004).
- Kaifu, Y., Aziz, F. & Baba, H. Hominid mandibular remains from Sangiran: 1952–1986 collection. *Am. J. Phys. Anthropol.* **128**, 497–519 (2005).
- Chang, C. H. et al. The first archaic *Homo* from Taiwan. *Nature Commun.* **6**, 6037 (2015).





**Extended Data Figure 1 | CT-based images of the SOA-MM4 mandible and photos of the SOA-MM6 incisor.** **a–i**, SOA-MM4 mandible. Surface-rendered images of superior (**a**), lateral (**b**), inferior (**c**), lingual (**d**), anterior (**e**), and posterior (**f**) views. Sagittal (**h**) and horizontal (**i**). CT sections at the plane indicated by the green (**g**) and red (**h**) lines. aMF, a branch of the mandibular foramen; ARR, anterior ramus root; LP, lateral prominence; M<sub>1</sub>, M<sub>1</sub> alveolus; M<sub>2</sub>, M<sub>2</sub> alveolus; M<sub>3</sub>, M<sub>3</sub> alveolus; MC, mandibular canal; Mas, line for the masseter muscle attachment; MHL, mylohyoid line; pbMC, posterior branch of the mandibular canal; SLT, superior lateral torus. **j–k**, SOA-MM6 mandibular incisor (I<sub>1/2</sub>) fragments. The crown (**j**, SOA-MM6a) and a root (**k**, SOA-MM6b) fragments were used for laser ablation uranium-series dating. The specimen was deposited before at least 0.55 Ma<sup>7</sup>. Note the bevelled occlusal wear surface (arrow). Scale bar, 5 mm.



**Extended Data Figure 2 | Linear metric comparisons of the mandibles and permanent teeth.** a–e, Scatter plots of the mandibular corporal dimensions (a, b) and permanent tooth crown diameters (c–e). We identify SOA-MM1 as M<sub>1</sub> (e), but there remains a slight possibility that this tooth is M<sub>2</sub> (f). Metric data of SOA-MM4: corpus height at M<sub>2</sub>, 18 mm; corpus height at M<sub>2</sub>/3, 18.5 mm; corpus width at M<sub>2</sub>, 12.5 mm; corpus width at M<sub>2</sub>/3, 13 mm.



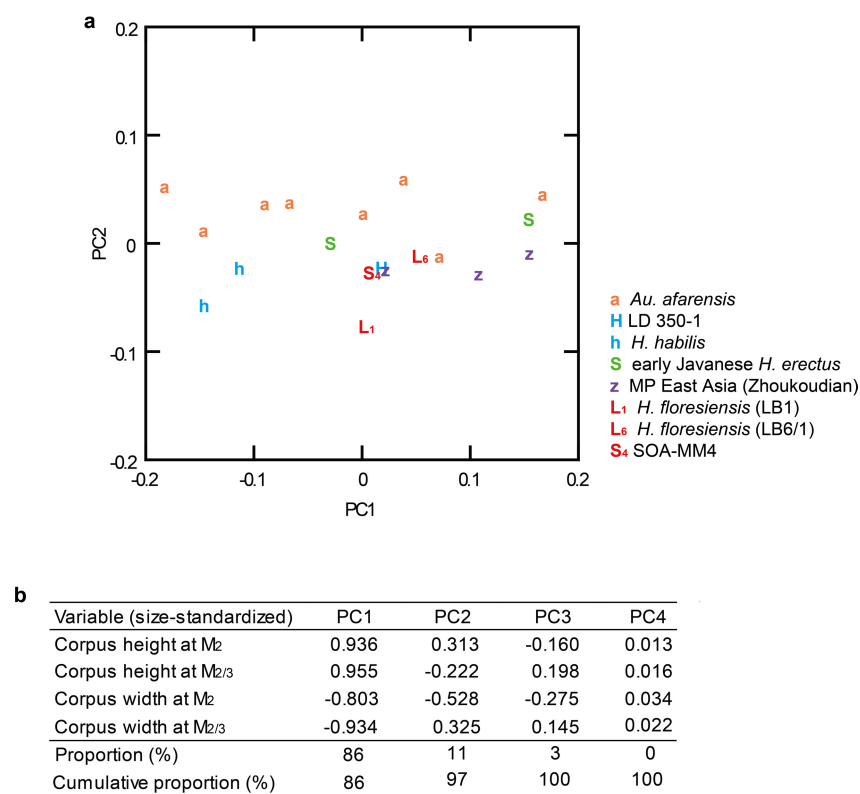
**Extended Data Figure 3 | Mandibular comparisons.** a–m, *H. habilis sensu lato*: OH 13 (a, late adolescent), OH 37 (b, horizontally flipped image), KNM-ER 1802 (c, late adolescent), KNM-ER 3734 (d, horizontally flipped image), KNM-ER 60000 (e, horizontally flipped image) (photo by F. Spoor, copyright National Museums of Kenya); early Javanese *H. erectus*, Sangiran 1b (f), Sangiran 9 (g), Sangiran 22 (h), Sb 8103 (i), Sangiran 21 (j); Liang Bua *H. floresiensis*: LB1 (k), LB6/1 (l, horizontally flipped image; the corpus is distorted); (m) SOA-MM4. Scale bar, 30 mm. Note that the

*H. habilis* mandibles tend to exhibit a thicker corpus, the position of the basal ramus (filled arrow) that is shifted laterally relative to the corpus midline axis, a prominent posterior part of the alveolar prominence (filled triangle), and a wider extramolar sulcus between the anterior ramus root (open arrow) and the molar row. The early Javanese *H. erectus* sample is variable but includes specimens with weaker expressions in these traits. The Liang Bua *H. floresiensis* and the SOA-MM4 mandibles share such derived features with early Javanese *H. erectus*.



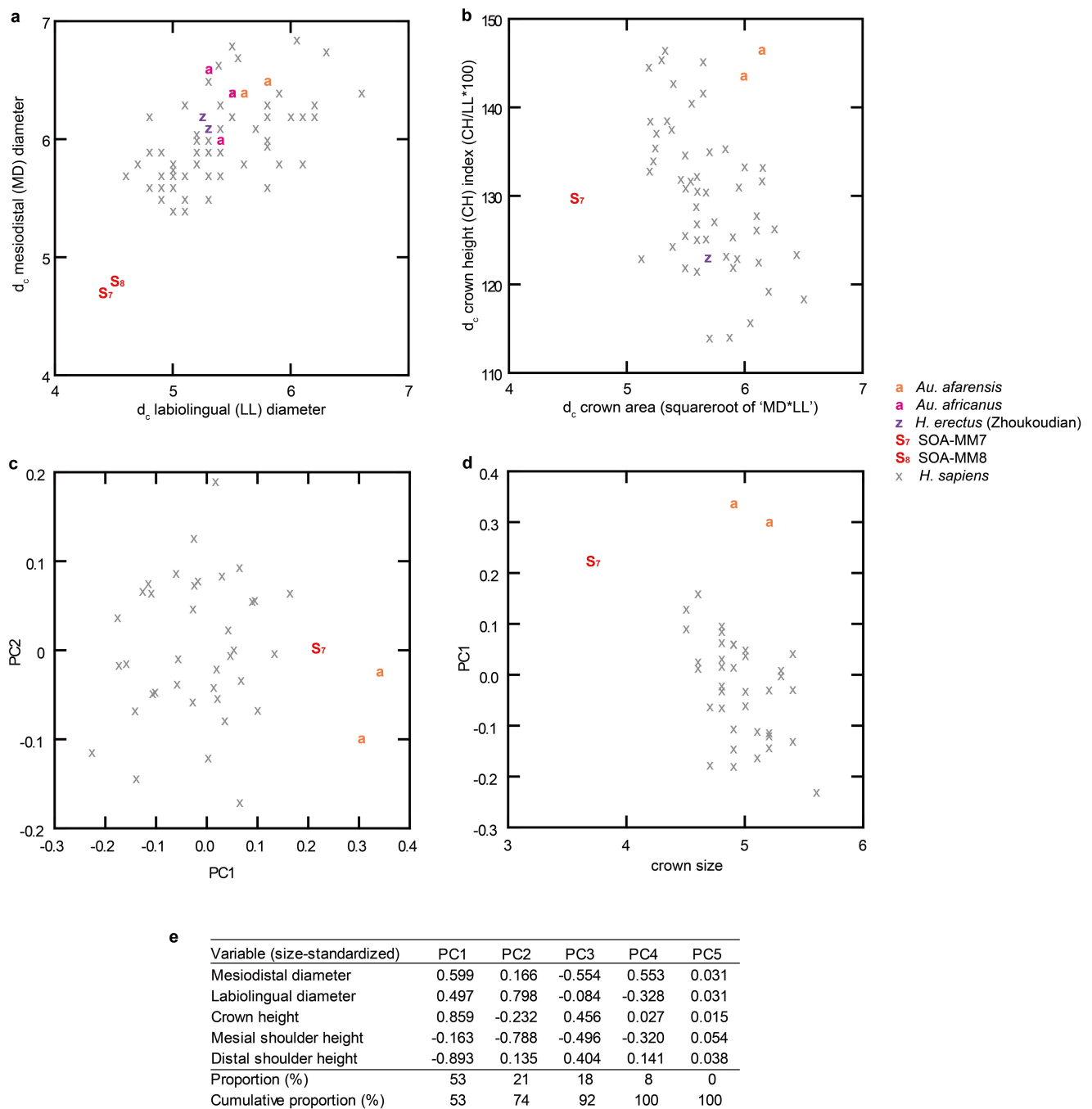


**Extended Data Figure 4 | Comparisons of the hominin mandibles and teeth from So'a Basin (Mata Menge) and *H. floresiensis* from Liang Bua.** **a**, SOA-MM4 mandible. **b**, **c**, Right lateral and left lateral (horizontally flipped) views of LB1. **d**, Right lateral view of LB6/1. **e**, **f**, The SOA-MM3 and LB1 P<sub>3</sub>s, respectively. **g**, SOA-MM1 M<sub>1</sub>. **h–j**, Occlusal views of SOA-MM4 (**h**), LB1 (**i**), and LB6/1 (**j**) mandibles. Scale bar, 10 mm.



**Extended Data Figure 5 | Principal component analyses of the four size-standardized mandibular measurements.** **a**, Scatter plot of the PC scores. **b**, Component loading of each PC. PC1 does not distinguish *Homo* from *Au. afarensis*, but *Au. afarensis* and post-*habilis* *Homo* are relatively

well-separated in PC2. SOA-MM4 belongs to the cluster of *Homo* in this PC. SOA-MM4 occupies the space in between the two Liang Bua *H. floresiensis* mandibles, suggesting their shared lateral corporal shape.



**Extended Data Figure 6 | Metric analyses of mandibular deciduous canines. a–e,** Comparisons of the crown length and breadth (a), and relative crown height (b). Results of the PCA based on size-adjusted five crown diameters (c, d) and the component loadings of each PC (e). ‘Crown size’ = geometric mean of the five crown diameters used.

*Au. afarensis* and *H. sapiens* are indistinguishable in crown size (d) but they are discriminated from each other by PC1 ( $P < 0.00002$ ,  $t$ -test). SOA-MM7 occupies an intermediate position between *Au. afarensis* and *H. sapiens*, suggesting its moderately primitive crown configuration. The other PCs did not discriminate *Au. afarensis* and *H. sapiens*.



**Extended Data Table 1 | 2014 Hominin fossil collection from Mata Menge**

Specimen No.	Catalogue No.	Date of discovery	Portion
SOA-MM1	MM14-T32D-F191	2014 Oct 08	broken crown of left M <sub>1</sub> (or M <sub>2</sub> )
SOA-MM2	MM14-T32C-F234	2014 Oct 14	complete crown and root of left I <sup>1</sup>
SOA-MM3	MM14-T32D-F384	2014 Oct 14	hominin cranial fragment?
SOA-MM4	MM14-T32C-F277	2014 Oct 14	right mandibular body
SOA-MM5	MM14-T32C-F452	2014 Oct 16	complete crown and root of right P <sup>3</sup>
SOA-MM6	MM14-T32B-F94	2014 Oct 18	broken root and crown fragments of right I <sub>1/2</sub>
SOA-MM7	MM14-T32C-dry sieve	2014 Oct 21	nearly complete left d <sub>c</sub>
SOA-MM8	MM14-T32B-dry sieve	2014 Oct 24	nearly complete right d <sub>c</sub>

All specimens are housed in the Geology Museum in Bandung.

**Extended Data Table 2 | Hominin teeth from Mata Menge as compared to those of Liang Bua *H. floresiensis*<sup>a</sup>**

Specimen	Tooth	Side	Wear <sup>b</sup>	Crown diam.				Cervical diam. <sup>e</sup>		Root length <sup>f</sup>
				MD <sup>c</sup>	MD <sup>c</sup>	BL <sup>c</sup>	Height <sup>d</sup>			
				as measured	corrected			MD	BL	
<b>SOA-MM2</b>	<b>I<sup>1</sup></b>	<b>L</b>	<b>5</b>	<b>7.6</b>	<b>8.0</b>	<b>6.4</b>	–	<b>5.5</b>	<b>5.7</b>	<b>12.0</b>
LB15/2	I <sup>1</sup>	L	7	–	–	≥ 6.2	–	5.2	6.2	12.0–13.0
<b>SOA-MM5</b>	<b>P<sup>3</sup></b>	<b>R</b>	<b>3</b>	<b>6.6</b>	<b>6.8</b>	<b>9.4</b>	–	<b>4.4</b>	<b>(8.4)</b>	<b>(12.5)</b>
LB1/1 <sup>i</sup>	P <sup>3</sup>	R & L	3.5	6.85	7.0	9.25	–	4.85	9.0	14.95
<b>SOA-MM6</b>	<b>I<sub>1/2</sub></b>	<b>R</b>	<b>5?</b>	–	–	–	–	–	–	–
<b>SOA-MM1</b>	<b>M<sub>1</sub> (or M<sub>2</sub>)</b>	<b>L</b>	<b>2</b>	<b>9.7</b>	<b>9.7</b>	<b>8.9</b>	–	–	–	–
LB1 <sup>i</sup>	M <sub>1</sub>	R & L	4.75	9.25	9.6	10.5	–	8.2	9.1	12.4, 12.45 <sup>g</sup>
LB6/1 <sup>i</sup>	M <sub>1</sub>	R & L	4.5	8.9	9.2	10.0	–	–	8.8	–
LB1 <sup>i</sup>	M <sub>2</sub>	R & L	3.75	9.8	10.1	10.2	–	8.8	8.9	12.5, 11.75 <sup>g</sup>
LB6/1 <sup>i</sup>	M <sub>2</sub>	R & L	3.5	9.45	9.7	9.55	–	–	–	–
<b>SOA-MM7</b>	<b>d<sub>c</sub></b>	<b>L</b>		<b>4.7</b>	<b>4.7</b>	<b>4.4</b>	<b>(5.8)</b>	<b>3.9</b>	<b>3.4</b>	–
<b>SOA-MM8</b>	<b>d<sub>c</sub></b>	<b>R</b>		<b>4.8</b>	<b>4.9</b>	<b>4.5</b>	–	<b>3.9</b>	<b>3.4</b>	–

<sup>a</sup>Measurements of the Liang Bua hominins cited from ref. 16.<sup>b</sup>Scored following ref. 49.<sup>c</sup>Measured following the method of ref. 39.<sup>d</sup>Buccal crown height.<sup>e</sup>Cervical diameters as defined by ref. 50.<sup>f</sup>Buccal root length.<sup>g</sup>Lengths for mesial and distal roots, respectively.<sup>i</sup>Average of the right and left sides.

Extended Data Table 3 | Comparative fossil samples

Sample	Age (Ma)	Portion	Composition	Data source
<i>H. floresiensis</i> (Liang Bua)	0.1-0.06	Mandible	LB1 <sup>#</sup> , 6/1 <sup>#</sup>	Originals
		Teeth	LB1*, 6/1*	Ref. 16
<i>Au. afarensis</i>	3.6-3.0	Mandible	A.L. 198-22 <sup>#</sup> , 225-8, 315-22, 330-5 <sup>#</sup> , 417-1a <sup>#</sup> , 436-1 <sup>#</sup> , 437-2, 438-1 <sup>#</sup> , 444-2, 620-1, 188-1 <sup>#</sup> , 198-1, 207-13, 266-1, 333w1a,b, 333w-32+60, MAL1/12 <sup>#</sup> , MAK1/2 <sup>#</sup> , LH4	Ref. 51
"Earliest <i>Homo</i> "	2.8	Mandible	LD350-1 <sup>#</sup>	Ref. 9
<i>H. habilis</i>	2.3-1.6	Mandible	KNM-ER 3734, 1805, 60000; OH 13 <sup>#</sup> , 37 <sup>#</sup>	Originals, Ref.40
		Teeth	A.L. 666-1; KNM-ER 1502*, 1507*, 1508*, 1590, 1801*, 1802*, 1805*, 1813; 2597*, 2601*, 60000; OH7*, 13*, 16*, 37*, 39; L7-279*, 628-10, 894-1; Omo 75-14*, 75s-15*, 195-1630*	Refs. 3,13, 40
Dmanisi <i>Homo</i>	1.75	Teeth	D211, 2600, 2700, 2735	Ref.15
Early Javanese <i>H. erectus</i> (older)	≥1.2-1.0	Mandible	Sangiran 1b <sup>#</sup> , 5, 9 <sup>#</sup> , 22	Originals, Ref.52
		Teeth	Sangiran 1b*, 4, 5*, 6b*, 6b, 7-35, 7-42*, 7-43*, 7-58, 7-61*, 7-62*, 7-63*, 7-64*, 7-65*, 7-76*, 7-78*, 7-84*, 7-85, 7-86, 22*; Bk 7905*	Ref. 3
Early Javanese <i>H. erectus</i> (younger)	1.0-0.8	Teeth	Sangiran 7-20*, 7-21*, 7-22*, 7-27, 7-31, 7-32; Sb 8103*, Ng 8503*	Ref. 3
Middle Pleistocene East Asian <i>Homo</i>	c. 0.75-0.05	Mandible	Lantian; Zhoukoudian G1 <sup>#</sup> , H1 <sup>#</sup> , K1 <sup>#</sup> , PA86; Penghu 1	Originals, casts, Ref.53
		Teeth	Zhoukoudian 1/2, 4, 19, 34, 35, 36/37, 43, 44, 45, 97, 98, 99, 107, 108, 137; Lantian PA102; Hexian PA834, 835, 838, 839, AN1644; Chaoxian; Xujiayao PA1480-1, 1480-3	Refs.3, 41, 42

#Specimens included in the PCAs.

\*Specimens included in the EFAs.

Refs. 51–53 are cited in this table.



Extended Data Table 4 | Comparative *Homo sapiens* dental sample

	Remarks	N <sup>a</sup>	Repository <sup>b</sup>
<b>Prehistoric Southeast Asia</b>			
Flores*	Aimere, Gua Alo, Gua Nempong, Liang Bua, Liang Momer, Liang Toge, Liang X	9	NBC, ARKENAS
Java*	Hoekgrot, Wajak	3	NBC
Malaysia*	Guar Kepah	19	NBC
Vietnam*	Mai Da Dieu, Mai Da Nuoc, Hang Chim, Dong Cang, Con Co Ngua	73	IAH
<b>Australia/Melanesia</b>			
New Guinea*		30	AMNH, MH
Indigenous Australian/Tasmanian*		19	AMNH
<b>Southeast Asia</b>			
Philippine Negrito*		20	MH
Others	Andaman, Indonesia, Malaysia, Nicobar, Philippine, Singapore, Thailand	57	AMNH, MH
<b>Northeast Asia</b>			
Northeast Asia	China, Chukuci, Korea, Mongol, Yukagir	18	AMNH
<b>Africa</b>			
Bushman		17	AMNH, MH
African Pygmy		20	MH
South Africa	Excluding Bushman	26	AMNH
East Africa		45	AMNH
West Africa	Excluding Pygmy	55	AMNH
<b>Indo/Europe</b>			
India		6	AMNH
German		65	AMNH
Others	Hungary, Poland, Sweden	8	AMNH
Total		490	

<sup>a</sup>Number of individuals.<sup>b</sup>NBC, Naturalis Biodiversity Center, Leiden; ARKENAS, National Research and Development Centre for Archaeology, Jakarta; AMNH, American Museum of Natural History, New York; MH, Musée de l'Homme, Paris; IAH, Institute of Archaeology, Hanoi.

\*Samples included in the EFAs.

# Age and context of the oldest known hominin fossils from Flores

Adam Brumm<sup>1,2\*</sup>, Gerrit D. van den Bergh<sup>3\*</sup>, Michael Storey<sup>4</sup>, Iwan Kurniawan<sup>5\*</sup>, Brent V. Alloway<sup>3,6</sup>, Ruly Setiawan<sup>3,7</sup>, Erick Setiyabudi<sup>5</sup>, Rainer Grün<sup>1,8</sup>, Mark W. Moore<sup>9</sup>, Dida Yurnaldi<sup>3,7</sup>, Mika R. Puspaningrum<sup>3</sup>, Unggul P. Wibowo<sup>3,5</sup>, Halmi Insani<sup>5</sup>, Indra Sutisna<sup>5</sup>, John A. Westgate<sup>10</sup>, Nick J. G. Pearce<sup>11</sup>, Mathieu Duval<sup>12</sup>, Hanneke J. M. Meijer<sup>13</sup>, Fachroel Aziz<sup>5</sup>, Thomas Sutikna<sup>3,14</sup>, Sander van der Kaars<sup>15,16</sup>, Stephanie Flude<sup>17</sup> & Michael J. Morwood<sup>3‡</sup>

**Recent excavations at the early Middle Pleistocene site of Mata Menge in the So'a Basin of central Flores, Indonesia, have yielded hominin fossils<sup>1</sup> attributed to a population ancestral to Late Pleistocene *Homo floresiensis*<sup>2</sup>. Here we describe the age and context of the Mata Menge hominin specimens and associated archaeological findings. The fluvial sandstone layer from which the *in situ* fossils were excavated in 2014 was deposited in a small valley stream around 700 thousand years ago, as indicated by <sup>40</sup>Ar/<sup>39</sup>Ar and fission track dates on stratigraphically bracketing volcanic ash and pyroclastic density current deposits, in combination with coupled uranium-series and electron spin resonance dating of fossil teeth. Palaeoenvironmental data indicate a relatively dry climate in the So'a Basin during the early Middle Pleistocene, while various lines of evidence suggest the hominins inhabited a savannah-like open grassland habitat with a wetland component. The hominin fossils occur alongside the remains of an insular fauna and a simple stone technology that is markedly similar to that associated with Late Pleistocene *H. floresiensis*.**

Mata Menge is located near the northwestern margin of the So'a Basin, a ~400 km<sup>2</sup> geological depression in central Flores (Fig. 1). The basement substrate consists of the Ola Kile Formation (OKF), a greater than 100-m-thick sequence of indurated volcanoclastic deposits<sup>3,4</sup>. Zircon fission-track (ZFT) age determinations date the upper OKF to 1.86 ± 0.12 million years ago (Ma) (ref. 4). The ~5° southward dipping volcanic breccias of the OKF are associated with a former volcanic centre, the Welas Caldera, on the northwestern margin of the basin (Fig. 1). The OKF is unconformably overlaid by the Ola Bula Formation (OBF)<sup>3,4</sup>. A focus of palaeoanthropological research since the 1950s<sup>5–14</sup>, the OBF is up to 120 m thick and comprises an intra-basinal fossil- and stone artefact-bearing sequence composed largely of undistorted volcanic, fluvial, and lacustrine sediments deposited between 1.8 and 0.5 Ma<sup>3,4</sup> (Supplementary Information Table 1). An extensive lacustrine sequence—the 'Gero Limestone Member' (GLM)—caps the basin infill and registers the formation of a basin-wide freshwater lake.

The total preserved thickness of the OBF at Mata Menge is 40 m (Fig. 1). The uppermost interval of the GLM, with a thickness of 9 m, crops out on a hill 600 m west (excavation number 35, or E-35). The two main fossil-bearing intervals at Mata Menge form part of a roughly NNW-SSE trending palaeovalley-fill sequence dominantly occupied by

cut-and-fill fluvial and clay-rich, water-supported mass flow deposits (mudflows). The studied upper fossiliferous interval, which contains the hominin fossils, is exposed at the head of a modern stream valley at the base of a hill (height = 397 m). This less than 30-cm-thick OBF sandstone, named Layer II, is well-consolidated, fine- to medium-grained, and contains locally faint parallel laminations in the lower part, as well as numerous water-worn volcanic pebbles (<60 mm). Layer II is discontinuous towards the west and east, and it has an irregular lower bedding plane that cuts down into the underlying unit, a well-developed, consolidated palaeosol (Layer III). A ~6.5-m-thick sequence of mudflow deposits (Layers I-a to I-f) overlies Layer II and is separated from it by a generally sharp contact surface. Layer II represents the deposit of a small, sinuous stream tributary with a NNW to SSE flow direction, as deduced from the slight decrease in elevation of the top of Layer II in the same direction (20 cm over a horizontal distance of 17 m).

We conducted a 50 m<sup>2</sup> excavation (E-32) into Layer II in 2013 (Fig. 1, Extended Data Figs 1 and 2), yielding fossils of *Stegodon florensis*<sup>9</sup>, giant rat (*Hooijeromys nusatenggara*, first described in ref. 15), Komodo dragon (*Varanus komodoensis*), and crocodile, as well as stone artefacts (Fig. 2). In 2014, we exposed Layer II over a larger area, recovering seven hominin fossils (six teeth and a mandible fragment), and an undiagnostic hominin cranial fragment<sup>1</sup>. The hominin fossils were embedded in the sandstone matrix of Layer II near the stratigraphic interface with the overlying mudflow deposit (Extended Data Fig. 2).

Layers I-a to I-f are clearly related to eruptive activity within the Welas Caldera, then occupied by a lake. Four articulated thoracic vertebrae of *S. florensis* were recovered from Layer II (Fig. 2k). These are the only articulated stegodont elements so far recovered at Mata Menge, indicating relatively limited post-mortem modification before burial by mudflows. We hypothesize that the artefacts and faunal remains, including hominin elements, were transported short distances by the stream that deposited Layer II, before mudflows originating from within the Welas Caldera inundated these valleys with metre-thick muddy debris. The presence of elements from multiple hominin individuals could be the result of the same volcanic event that triggered the mudflows. Presently, however, it is not possible to estimate the time interval separating the deaths of the hominins from the deposition of the mudflows.

<sup>1</sup>Research Centre of Human Evolution, Environmental Futures Research Institute, Griffith University, Nathan, Queensland 4111, Australia. <sup>2</sup>School of Earth & Environmental Sciences, University of Wollongong, Wollongong, New South Wales 2522, Australia. <sup>3</sup>Centre for Archaeological Science, School of Earth & Environmental Sciences, University of Wollongong, Wollongong, New South Wales 2522, Australia. <sup>4</sup>QuaDlab, Natural History Museum of Denmark, University of Copenhagen, DK-1350 Copenhagen, Denmark. <sup>5</sup>Geology Museum, Bandung 40122, Indonesia. <sup>6</sup>School of Geography, Environment and Earth Sciences, Victoria University, Wellington 6012, New Zealand. <sup>7</sup>Center for Geological Survey, Geological Agency, Bandung 40122, Indonesia. <sup>8</sup>Research School of Earth Sciences, The Australian National University, Canberra, Australian Capital Territory 2601, Australia. <sup>9</sup>Stone Tools and Cognition Hub, Archaeology, University of New England, Armidale, New South Wales 2351, Australia. <sup>10</sup>Department of Earth Sciences, University of Toronto, Toronto, Ontario M5S 3B1, Canada. <sup>11</sup>Department of Geography & Earth Sciences, Aberystwyth University, Aberystwyth SY23 3DB, UK. <sup>12</sup>Geochronology, Centro Nacional de Investigación sobre la Evolución Humana (CENIEH), Paseo de Atapuerca, 3, 09002-Burgos, Spain. <sup>13</sup>University Museum of Bergen, University of Bergen, 5007 Bergen, Norway. <sup>14</sup>Pusat Penelitian Arkeologi Nasional (ARKENAS), Jakarta 12510, Indonesia. <sup>15</sup>Cluster Earth & Climate, Faculty of Earth and Life Sciences, Vrije Universiteit, 1081 HV Amsterdam, The Netherlands. <sup>16</sup>School of Earth, Atmosphere and Environment, Monash University, Clayton, Victoria 3800, Australia. <sup>17</sup>School of Geosciences, The University of Edinburgh, Edinburgh EH8 9AD, UK.

\*These authors contributed equally to this work.

‡Deceased.

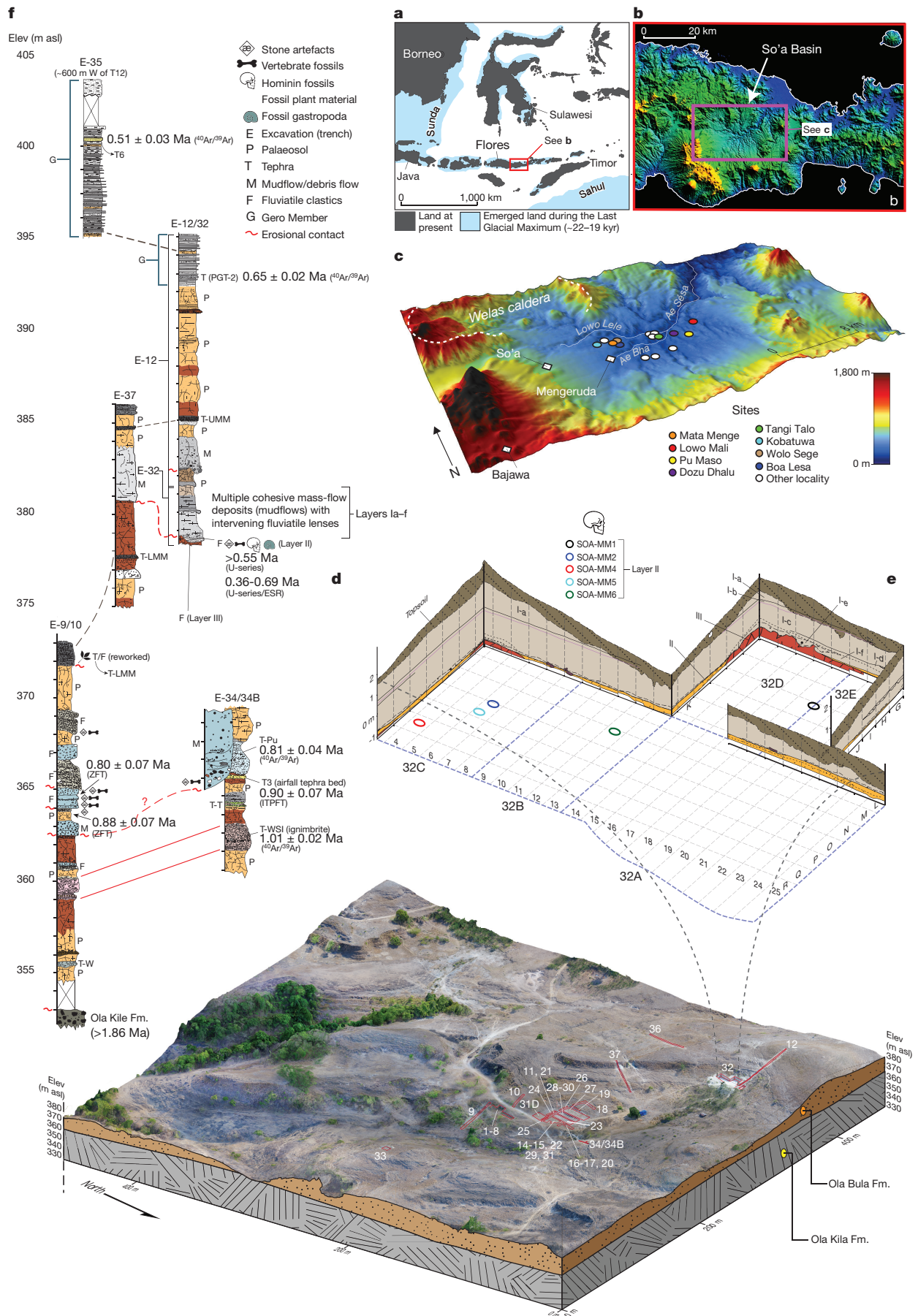


Figure 1 | See next page for caption.

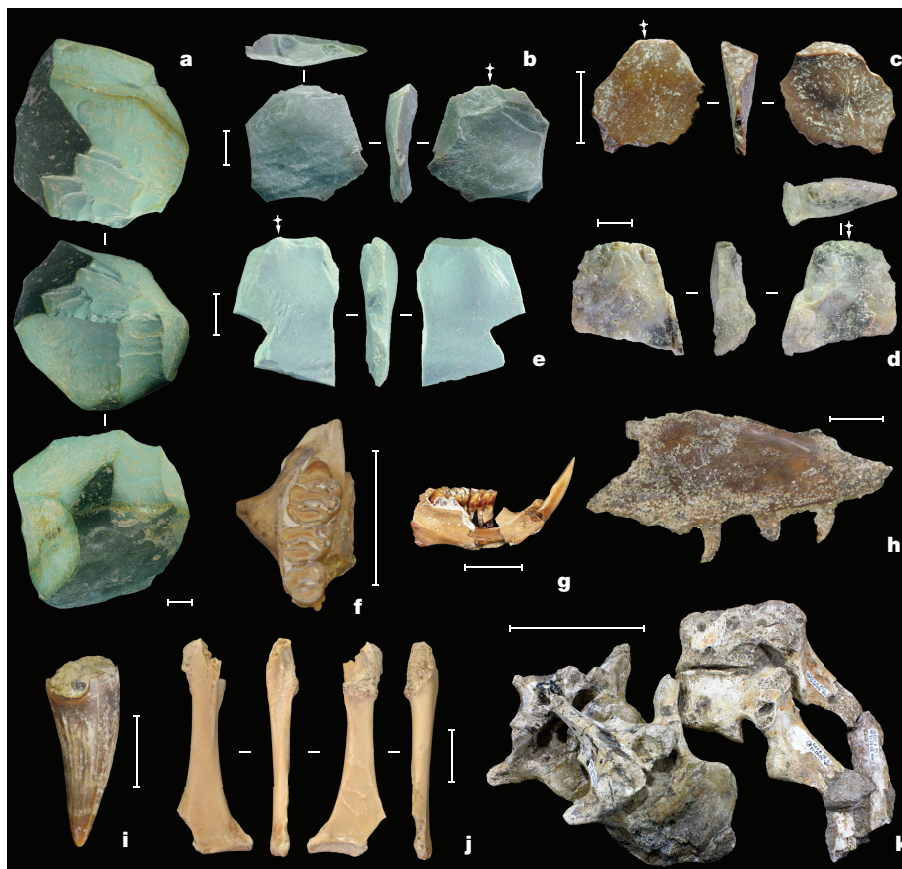


**Figure 1 | Context and chronology of the hominin fossils at Mata Menge.** **a, b**, location of Flores and the So'a Basin. **c**, Digital elevation map of the So'a Basin, with location of Mata Menge and other sites mentioned in the text. A single outlet of the main river system (the Ae Sésa) drains the basin via a steep-walled valley towards the northeast. **d**, Stratigraphy and chronology of the main fossil-bearing intervals and intervening Ola Bula Formation (OBF) deposits at Mata Menge. Several basin-wide key marker tephra beds that are exposed in the hill flank on the northern side of Mata Menge (trench E-34/34B) are eroded in the central part of the stream valley, where they are replaced by a 4–5-m-thick sequence of tuffaceous mudflows with intervening fluvial lenses forming the lower fossil-bearing palaeovalley-fill sequence. **e, f**, Context of the hominin fossils; **f** is a 3D image of Mata Menge and surrounds, with excavated trenches outlined in red and labelled, and **e** is a 3D representation of the stratigraphy exposed by trench E-32A to E, with coloured ovals denoting the positions of *in situ* hominin fossils (SOA-MM1, 2 and 4–6) excavated from the fluvial sandstone unit, Layer II. The remaining hominin specimens were retrieved in the sieves. Trenches E-1 to E-8 were

excavated between 2004 and 2006, at the section originally excavated by Th. Verhoeven in the 1950s<sup>5,6</sup>. The remaining trenches were excavated between 2010 and 2015. E-12 is a slot-trench excavated into the side of a hill, revealing an 18-m-thick sequence of lacustrine clays and micritic limestones, fluvial sandstone beds, massive tuffaceous mudflows, well-developed clay-textured palaeosols, and numerous centimetre-thick basaltic tephra inter-beds from the middle-upper part of the OBF. At the base of this slot-trench, a less than 30-cm-thick fossil-bearing sandstone unit (Layer II)—from which all the hominin fossils were retrieved—was exposed. Tephra codes in **d** are as follows (top to bottom): T6 (upper inter-regional tephra); PGT-2 (Piga Tephra 2); T-UMM (Upper Mata Menge Tephra); T-LMM (Lower Mata Menge Tephra); T-Pu (Pu Maso Tephra); T3 (lower inter-regional tephra); T-T (Turakeo Tephra); T-WSI (Wolo Sege Ignimbrite); and T-W (Wolowawu Tephra). The original published  $^{40}\text{Ar}/^{39}\text{Ar}$  age for T-WSI is  $1.02 \pm 0.02$  Ma (ref. 14); however, when recalculated to the recently determined value for the age standard ACS-2 used in this study (1.185 Ma; see reference 25 in the Supplementary Information), T-WSI becomes  $1.01 \pm 0.02$  Ma.

A total of four new radioisotopic age determinations, with ages in sequential order and in accordance with the stratigraphic sequence, provide a chronology for the hominin fossils (Fig. 1; Supplementary Information). Near the base of the OBF at Mata Menge, a widespread ignimbritic marker bed (the Wolo Sege Ignimbrite; T-WSI) with an  $^{40}\text{Ar}/^{39}\text{Ar}$  age of  $1.01 \pm 0.02$  Ma (ref. 14; Fig. 1) is recognized on the combined basis of its stratigraphic association, unique depositional architecture, and glass-shard major element chemistry (Extended Data Fig. 3). In addition, the hominin find-locality in E-32 is situated 12.5 m stratigraphically above a previously reported ZFT date of  $0.80 \pm 0.07$  Ma from Mata Menge<sup>4</sup>. To verify this prior

estimate, we conducted isothermal plateau fission-track (ITPFT) dating of glass shards from an inter-regional tephra marker (T3) identified at several So'a Basin localities, including just above the T-WSI at Mata Menge (in E-34/34B), returning a weighted mean age of  $0.90 \pm 0.07$  Ma (based on two independent age determinations) (Supplementary Information Table 2).  $^{40}\text{Ar}/^{39}\text{Ar}$  single crystal dating of hornblende from the Pu Maso Ignimbrite (T-Pu) located just above T3 in E-34/34B yielded a weighted mean age of  $0.81 \pm 0.04$  Ma (Extended Data Fig. 4), which is stratigraphically consistent with that of underlying T3. These ages demonstrate that Layer II was deposited after  $\sim 0.80$  Ma.



**Figure 2 | Stone artefacts and fossils from Mata Menge.** All specimens are from the hominin fossil find-locality (Layer II fluvialite sandstone, trench E-32). **a**, Bifacial core (chlorite). **b, c**, Chert flakes. **d**, Chalcedony flake. **e**, Rhyolite flake. **f**, Right maxilla fragment (M1-M3), *Hooijeromys nusatenggara*. **g**, Left mandible fragment (m1-m3), *H. nusatenggara*.

**h**, Right maxilla fragment, *Varanus komodoensis*. **i**, Crocodile tooth. **j**, Right coracoid of a duck (cf. *Tadorna*). **k**, *Stegodon florensis* thoracic vertebrae in articulation (still partially embedded in sandstone matrix). Scale bars, 10 mm (**a–j**); 100 mm (**k**).

To further constrain the age of the hominin fossils, we carried out  $^{40}\text{Ar}/^{39}\text{Ar}$  dating on one basaltic tephra and one rhyolitic tephra from the GLM above Layer II (E-12 and E-35). The GLM contains at least 85 crystal-rich tephra inter-beds of basaltic composition, collectively named the Piga Tephra (the lower 56 tephra are sequentially numbered PGT-1 to PGT-56). At Mata Menge, PGT-2 occurs 13.5 m above Layer II, and produced a  $^{40}\text{Ar}/^{39}\text{Ar}$  weighted mean age of  $0.65 \pm 0.02$  Ma from single crystal dating of hornblende (Extended Data Fig. 5). This is in accordance with the published ZFT age of a basaltic tephra inter-bed from the lower part of the GLM ( $0.65 \pm 0.06$  Ma)<sup>4</sup>. Finally, a biotite-bearing vitric-rich ash of distinctive rhyolitic composition (T6; Extended Data Fig. 3) from the top of the GLM has an  $^{40}\text{Ar}/^{39}\text{Ar}$  age of  $0.51 \pm 0.03$  Ma, based on the weighted mean of single grain feldspar analyses (Extended Data Fig. 6). Thus, the hominin fossils constrained by the lowermost of these two radioisotopic dates within the GLM have a minimum age of  $\sim 0.65$  Ma.

We also conducted uranium series (U-series) dating of a hominin tooth root fragment (SOA-MM6) from Layer II, and combined U-series and electron spin resonance (ESR) dating of two *S. florensis* molars excavated *in situ* from the same sedimentary context (Extended Data Fig. 7; Supplementary Information). U-series dating of the hominin tooth root independently confirms that this specimen has an age of at least 0.55 Ma, whereas combined U-series/ESR dating indicates minimum and maximum ages of around 0.36 Ma and 0.69 Ma, respectively, for the *Stegodon* molars. In sum, therefore, we have used multiple dating methods to establish an age of  $\sim 0.70$  Ma for the hominin fossils.

Our systematic, high volume excavations ( $\sim 560$  m<sup>2</sup>) at Mata Menge between 2010 and 2015 yielded many fossil vertebrate remains (Supplementary Information). To date, 75% of the  $>7,000$  vertebrate fossils recovered from E-32 have been analysed, and include *S. florensis* (23.7% of the number of identified specimens (NISP)), *V. komodoensis* (0.6% of NISP), freshwater crocodiles (3.7% of NISP), frogs (0.3% of NISP), murine rodents (15.6% of NISP), and birds (0.5% of NISP), the remainder comprising unidentifiable bone fragments. From the lower fossil-bearing interval (E-1 to 8 and E-11 to 31D), the remains of least 120 *S. florensis* individuals are represented by dental elements spanning all ontogenetic stages<sup>16</sup>. The age profile of this death-assemblage corresponds to that of a living population. The lack of age-selective mortality fits a mass-death event, unlike the juvenile-dominated pattern encountered in the *Stegodon* death-assemblage of the *H. floresiensis* type-locality, Liang Bua<sup>17</sup>. In Layer II, remains of juvenile, sub-adult, intermediate-aged, and old *Stegodon* individuals are also present, but the minimum number of individuals (MNI = 15) is too low to allow construction of a reliable age profile.

We conducted carbon and oxygen isotope analysis of tooth enamel samples collected from several *S. florensis* and murine rodent individuals from the two fossil-bearing levels at Mata Menge (Extended Data Fig. 8). The results indicate a diet dominated by C<sub>4</sub> grasses, suggesting both animals were grazers, and implying that open grasslands were the major vegetation type in the So'a Basin. The recovery of rare fossils of rails, swans, ducks, eagles, and eagle owls from the lower trenches ( $\sim 0.80$  to  $0.88$  Ma) provides further evidence for the presence of a savannah-like biome with wetland habitats, as well as scattered patches of forest<sup>18</sup>. Fossil pollen and phytoliths from both fossil levels offer additional indications that grasses dominated the early Middle Pleistocene vegetation (Supplementary Information Table 9). Abundant moulds and casts of two freshwater gastropod species (*Cerithiidea*) were recovered from Layer II, pointing to the existence of permanent freshwater bodies in the ancestral stream valley.

Our excavations uncovered 149 *in situ* stone artefacts in E-32, including 47 artefacts from Layer II, in direct association with the hominin remains (Fig. 2; Extended Data Fig. 9). Some of the artefacts from E-32 are lightly to heavily abraded from low-energy water transport<sup>19</sup>, but 74.5% are in fresh, as-struck condition, suggesting minimal dislocation from nearby stone-flaking areas. Overall, the E-32 assemblage reflects a technologically straightforward core-and-flake approach to

stoneworking<sup>20</sup>. As yet, no butchery marks have been identified on the faunal remains at Mata Menge.

Notably, the tools and flaking technology in E-32 are nearly identical in size and nature, respectively, to the assemblage dating some 110 thousand years (kyr) earlier at Mata Menge<sup>13,21–23</sup>, including 1,186 analysed stone artefacts from E-23 and E-27 excavated between 2011 and 2014 (Supplementary Information Table 6). The E-32 assemblage is also technologically similar to the artefacts from Liang Bua, dating  $\sim 600$  kyr later<sup>13,24</sup> and associated with *H. floresiensis*<sup>25,26</sup>. The long persistence of this technology<sup>13</sup> suggests stability in the behaviour of *H. floresiensis*<sup>1</sup>. In contrast, the only lithic assemblage recovered *in situ* below the T-WSI—which has a minimum age of  $1.01 \pm 0.02$  Ma and is the oldest known technology from Flores<sup>14</sup>—while similar, features a typologically distinct element: large Acheulean pick-like implements<sup>27</sup> associated elsewhere with cognitively advanced tool-making<sup>20,28–30</sup>. It is unclear why these artefacts are absent from the later technology of Flores. A shift to more arid conditions could have stimulated a series of technological changes. Alternatively, the earliest inhabitants of Flores may have responded to the limited resources of the island by reducing the complexity of their tool-making repertoire to the minimum required for survival.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 30 December 2015; accepted 11 March 2016.**

- van den Bergh, G. D. *et al.* *Homo floresiensis*-like hominin fossils from the early Middle Pleistocene of Flores. *Nature* <http://dx.doi.org/10.1038/nature17999> (2016).
- Brown, P. *et al.* A new small-bodied hominin from the Late Pleistocene of Flores, Indonesia. *Nature* **431**, 1055–1061 (2004).
- Morwood, M. J., O'Sullivan, P. B., Aziz, F. & Raza, A. Fission-track ages of stone tools and fossils on the east Indonesian island of Flores. *Nature* **392**, 173–176 (1998).
- O'Sullivan, P. B. *et al.* Archaeological implications of the geology and chronology of the Soa Basin, Flores, Indonesia. *Geology* **29**, 607–610 (2001).
- Maringer, J. & Verhoeven, Th. Die Steinartefakte aus der *Stegodon*-Fossilschicht von Mengeruda auf Flores, Indonesien. *Anthropos* **65**, 229–247 (1970).
- Maringer, J. & Verhoeven, Th. Die Oberflächenfunde aus dem Fossilgebiet von Mengeruda und Olabula auf Flores, Indonesien. *Anthropos* **65**, 530–546 (1970).
- Sondaar, P. Y. *et al.* Middle Pleistocene faunal turn-over and colonisation of Flores (Indonesia) by *Homo erectus*. *C. R. Acad. Sci.* **319**, 1255–1262 (1994).
- Morwood, M. J. *et al.* Stone artefacts from the 1994 excavation at Mata Menge, West Central Flores, Indonesia. *Aust. Archaeol.* **44**, 26–34 (1997).
- van den Bergh, G. D. *The Late Neogene Elephantoid-Bearing Faunas of Indonesia and their Palaeozoogeographic Implications. A Study of the Terrestrial Faunal Succession of Sulawesi, Flores and Java, including Evidence for Early Hominid Dispersal East of Wallace's Line* (Scripta Geologica 117, Nationaal Natuurhistorisch Museum, 1997).
- van den Bergh, G. D. *et al.* Did *Homo erectus* reach the island of Flores? *Bull. Indo. Pac. Pre. Hi.* **14**, 27–36 (1996).
- Morwood, M. J. *et al.* Archaeological and palaeontological research in central Flores, east Indonesia: results of fieldwork 1997–98. *Antiquity* **73**, 273–286 (1999).
- Aziz, F. & Morwood, M. J. in *Pleistocene Geology, Palaeontology and Archaeology of the Soa Basin, Central Flores, Indonesia* (eds Aziz, F., Morwood, M. J. & van den Bergh, G. D.) 1–18 (Spec. Publ. 36, Geological Survey Institute, 2009).
- Brumm, A. *et al.* Early stone technology on Flores and its implications for *Homo floresiensis*. *Nature* **441**, 624–628 (2006).
- Brumm, A. *et al.* Hominins on Flores, Indonesia, by one million years ago. *Nature* **464**, 748–752 (2010).
- Musser, G. G. The giant rat of Flores and its relatives east of Borneo and Bali. *Bull. Am. Mus. Nat. Hist.* **169**, 67–176 (1981).
- van den Bergh, G. D. *et al.* Taphonomy of *Stegodon florensis* remains from the early Middle Pleistocene archaeological site Mata Menge, Flores, Indonesia. Abstract book of the VIth International Conference on Mammoths and their relatives. S.A.S.G., Special Volume **102**, 207–208 (2014).
- van den Bergh, G. D. *et al.* The youngest *Stegodon* remains in Southeast Asia from the Late Pleistocene archaeological site Liang Bua, Flores, Indonesia. *Quat. Int.* **182**, 16–48 (2008).
- Meijer, H. J. M. *et al.* Avian remains from the Early/Middle Pleistocene of the So'a Basin, central Flores, Indonesia, and their palaeoenvironmental significance. *Palaeogeogr. Palaeoclimatol.* **440**, 161–171 (2015).

19. Shea, J. J. Artifact abrasion, fluvial processes, and “living floors” from the Early Paleolithic site of ‘Ubeidiya (Jordan Valley, Israel). *Geoarchaeology* **14**, 191–207 (1999).
20. Moore, M. W. The design space of stone flaking: implications for cognitive evolution. *World Archaeol.* **43**, 702–715 (2011).
21. Brumm, A. *et al.* Stone technology at the Middle Pleistocene site of Mata Menge, Flores, Indonesia. *J. Arch. Sci.* **37**, 451–473 (2010).
22. Moore, M. W. & Brumm, A. Stone artifacts and hominins in island Southeast Asia: new insights from Flores, eastern Indonesia. *J. Hum. Evol.* **52**, 85–102 (2007).
23. Moore, M. W. & Brumm, A. *Homo floresiensis* and the African Oldowan in *Interdisciplinary Approaches to the Oldowan* (eds Hovers, E. & Braun, D. R.) 61–69 (Springer, 2009).
24. Moore, M. W. *et al.* Continuities in stone flaking technology at Liang Bua, Flores, Indonesia. *J. Hum. Evol.* **57**, 503–526 (2009).
25. Morwood, M. J. *et al.* Archaeology and age of a new hominin from Flores in eastern Indonesia. *Nature* **431**, 1087–1091 (2004).
26. Sutikna, T. *et al.* Revised stratigraphy and chronology for *Homo floresiensis* at Liang Bua in Indonesia. *Nature* **532**, 366–369 (2016).
27. Brumm, A. & Moore, M. W. Biface distributions and the Movius Line: a Southeast Asian perspective. *Aust. Archaeol.* **74**, 32–46 (2012).
28. Beyene, Y. *et al.* The characteristics and chronology of the earliest Acheulean at Konso, Ethiopia. *Proc. Natl Acad. Sci. USA* **110**, 1584–1591 (2013).
29. Wynn, T. Archaeology and cognitive evolution. *Behav. Brain Sci.* **25**, 389–402 (2002).
30. Gowlett, J. A. J. The elements of design form in Acheulian bifaces: modes, modalities, rules and language, in *Axe Age: Acheulian Tool-making From Quarry to Discard* (eds Goren-Inbar, N. & Sharon, G.) 203–221 (Equinox, 2006).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** The So’a Basin project was funded by an Australian Research Council (ARC) Discovery grant (DP1093342) awarded to M.J.M. and A.B., and directed by M.J.M. (2010–2013) and G.v.d.B. (2013–2015). The Geological Survey Institute (GSI) of Bandung, Indonesia, provided additional financial and technical support. G.v.d.B.’s research was also supported by ARC Future Fellowship FT100100384. M.W.M. was funded by ARC grant DP1096558. Quadlab is funded by a grant to M.S. from the Villum Foundation. M.D. received funding from a Marie Curie International Outgoing Fellowship of the EU’s Seventh Framework Programme (FP7/2007–2013), awarded under REA Grant Agreement No. P10F-GA-2013-626474. B.V.A. received funding from a Victoria University of Wellington Science Faculty Research Grant (201255). For permission to undertake this research, we thank the Indonesian

State Ministry of Research and Technology (RISTEK), the former Heads of the Geological Agency (R. Sukiyar and Surono), the successive directors of the GSI (S. Permanandewi, Y. Kusumahbrata (formerly) and A. Pribadi) and Bandung’s Geology Museum (S. Baskoro and O. Abdurahman). Local research permissions were issued by the provincial government of East Nusa Tenggara at Kupang, and the Ngada and Nage Keo administrations. We also thank the Ngada Tourism and Culture and Education Departments for their ongoing support. In addition, we acknowledge support and advice provided by I. Setiadi, D. Pribadi, and Suyono (GSI), the Pusat Penelitian Arkeologi Nasional (ARKENAS) in Jakarta, and J. T. Solo of the provincial Culture and Tourism office in Kupang. Scientific and technical personnel involved in the fieldwork included: T. Suryana, S. Sonjaya, H. Oktariana, I. Sutisna, A. Rahman, S. Bronto, E. Sukandar, A. Gunawan, Widji, A. T. Hascaryo, Jatmiko, S. Wasisto, R. A. Due, S. Hayes, Y. Perston, B. Pillans, K. Grant, M. Marsh, D. McGahan, A. M. Saiful, B. Burhan, L. Siagian, D. Susanti, P. D. Moi, M. Tocheri, A. R. Chivas, and A. Cahyana. F. Wesselingh identified gastropod remains. Sidarto (GSI) provided digital elevation model data used in Fig. 1b. Geodetic surveys and measurements were conducted by E. E. Laksmana, A. Rahmadi, Y. Sofyan, and G. Hazell. J. Noblett constructed the Mata Menge 3D model, based on drone aerial photographs taken by K. Riza, T. P. Ertanto, and M. Faizal. The research team was supported by ~100 excavators and support personnel from the Ngada and Nage Keo districts. We thank L. Kinsley, Research School of Earth Sciences, The Australian National University, for assistance with mass spectrometric measurements.

**Author Contributions** A.B., G.D.v.d.B., I.K. and M.J.M. directed the Mata Menge excavations. M.S., B.V.A. and R.S. collected tephra samples and M.S. undertook  $^{40}\text{Ar}/^{39}\text{Ar}$  dating. G.D.v.d.B. described the site stratigraphy, with R.S., D.Y., S.F. and B.V.A. ITPFT-dating of T3 was jointly conducted by J.A.W. and B.V.A., while EMP analyses of all So’a Basin tephra were conducted by B.V.A. and R.S. Comparative trace element analyses of interregional tephra markers were jointly undertaken by J.A.W., N.J.G.P. and B.V.A. E.S., F.A. and T.S. oversaw key aspects of the field project. M.W.M. analysed the stone assemblage, and G.D.v.d.B., H.I., I.S., M.R.P., U.P.W. and H.J.M.M. analysed the fauna. M.R.P. conducted isotopic analyses, R.G. and M.D. undertook U/Th and ESR analyses of faunal remains, and S.v.d.K. carried out the palynological analysis. A.B. and G.D.v.d.B. prepared the manuscript, with contributions from other authors.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.D.v.d.B. ([gert@uow.edu.au](mailto:gert@uow.edu.au)) or A.B. ([a.brumm@griffith.edu.au](mailto:a.brumm@griffith.edu.au)).



## METHODS

For isothermal plateau fission track (ITPFT) dating of hydrated glass shards<sup>31</sup>, the population-subtraction technique was applied<sup>32,33</sup>. This technique is grain specific, in that every grain is examined separately under the microscope and a mean age is derived from a large number of shards. In this case, contamination can be readily monitored, and, if necessary, checked by electron microprobe analysis (EMA) of individual shards. Chemical homogeneity indicates derivation from a single eruptive event with a strong likelihood of uniform U content. Chunky, low vesicular glass shards in the size range of 500–250 µm were separated in order to maximise the surface area of the glass in the polished section and to optimise fission-track counting. ITPFT ages for T3 at two So'a Basin localities (Kopowatu, UT2382 and Lowo Mali, UT2383) are given in Supplementary Information Table 2, as well as operating conditions and the ages determined on the Huckleberry Ridge tephra internal standard. The single-crystal (sanidine) laser-fusion <sup>40</sup>Ar/<sup>39</sup>Ar age of Huckleberry Ridge tephra is  $2.003 \pm 0.014$  Ma ( $2\sigma$  error)<sup>34</sup>, and is indistinguishable from the PTF-corrected age of  $2.08 \pm 0.21$  Ma determined on UT1366 using the diameter-corrected procedure (DCFT)<sup>35</sup> (see Supplementary Information Table 2). All samples were irradiated at the same time in a single can. Ages were calculated using the zeta approach and  $\lambda_D = 1.551 \times 10^{-10}$  per year. Zeta value is  $301 \pm 3$  based on 6 irradiations, using the NIST SRM 612 glass dosimeter and the Moldavite tektite glass (Lhenice locality) with an <sup>40</sup>Ar/<sup>39</sup>Ar plateau age of  $14.34 \pm 0.08$  Ma (refs 36,37). Ages are those corrected for partial track fading (PTF), achieved by the isothermal plateau method (ITPFT)<sup>31</sup> and the diameter-corrected procedure (DCFT)<sup>35</sup>. Following irradiation, both T3 samples were subjected to a single heat treatment of 150°C for 30 days. After heating, the spontaneous and induced sample slides were simultaneously etched in 24% HF for 110 s. The Ds and Di track diameters were then measured and the Ds/Di ratio determined. Provided that the Ds/Di ratio was close to unity and that the samples were adequately etched — as evidenced by the average track diameters being within the range of 6–8 µm (ref. 31) — the samples were corrected for partial track fading (PTF) and the track densities determined. Area was then estimated using the point-counting method. The corresponding  $\rho_s/\rho_i$  ratio is equivalent to the true track density ratio. The age calculated from this ratio is therefore equivalent to the true age of the sample. The precision of ITPFT ages for some samples was improved by independent determinations made by different operators, in some cases using slightly different etching conditions. The weighted mean age and error of corrected ages for UT2382 and UT2383 is  $0.90 \pm 0.07$  Ma (see Supplementary Table 2).

**Glass shard major element determinations.** Glass shard major element determinations were conducted on all rhyolitic pyroclastic density current (PDC) and airfall deposits at Mata Menge, as well as potential correlatives from other So'a Basin sites. Glass shard major element data was acquired with a JEOL Superprobe (JXA-8230), using the ZAF correction method. Analyses were performed with 15 kV accelerating voltage, 8 nA beam current, and an electron beam defocused to between 20 to 10 µm. Standardization was achieved by means of mineral and glass standards. A rhyolitic glass standard (ATHO-G) was routinely used to monitor calibration in all analytical runs, and used to evaluate any day-to-day differences in the calibration. The large number of samples precluded conducting all analyses in a single batch. All analyses are normalized to 100 wt. % anhydrous, with H<sub>2</sub>O by difference being given, and total Fe is reported as FeO. Glass shard major element analyses are presented in Supplementary Information Table 3. Trace element analyses were conducted on individual glass shards from two distinct rhyolitic tephra marker beds of presumed distal source (T3 correlatives from Mata Menge, Lowo Mali and Kopowatu, and T6 from Mata Menge, respectively). This trace element data was then directly compared with reference data from potential distal tephra correlatives (that is, Youngest Toba Tuff (YTT), Middle Toba Tuff (MTT), Oldest Toba Tuff (OTT) and Unit E from ODP-758) acquired on the same instrument using the same standards and under the same analytical conditions<sup>38,39</sup> (see Extended Data Fig. 3i–m). Trace element analyses were performed by laser ablation (LA) ICP-MS, using a Coherent GeoLas ArF 193 nm Excimer LA system coupled to a Thermo Finnegan Element 2 sector field ICP-MS. Trace element data were collected for individual shards, with the majority of analyses performed using 20 µm ablation craters. Laser fluence was  $10 \text{ J cm}^{-2}$  at a repetition rate of 5 Hz for a 24 s acquisition. The minor <sup>29</sup>Si isotope was used as the internal standard, with SiO<sub>2</sub> (determined by EMPA) used to calibrate each analysis, after normalization to an anhydrous basis. The NIST 612 reference glass was used for calibration, taking concentrations from established sources<sup>40</sup>. A fractionation factor was applied to the data to account for analytical bias related to the different matrices of the reference standard and the sample. Explication of this factor as well as ICP-MS and laser operating conditions is given elsewhere<sup>41</sup>. The MPI-DING reference glass ATHO-G (ref. 42) was analysed as an unknown under the same operating conditions at the same time. Analytical precision is typically between  $\pm 5$ –10%, and accuracy is typically around  $\pm 5\%$ , when compared with the published GeoReM

concentrations for ATHO-G. Glass shard trace element analyses are presented in Supplementary Information Table 4.

**Single crystal laser fusion <sup>40</sup>Ar/<sup>39</sup>Ar dating.** We conducted single crystal laser fusion <sup>40</sup>Ar/<sup>39</sup>Ar dating of the Mata Menge volcanic units. Hornblende crystals (<2 mm in length) from the basaltic PGT-2 tephra sample (T Xll 252–261) were pre-concentrated along with other ferromagnesian minerals using standard heavy liquid techniques and then distinguished from pyroxene using a Bruker micro-xrf, followed by handpicking of individual grains under a binocular microscope. For the other samples, hornblende (1–2 mm in length from FLO15–15) and small feldspar crystals (<0.5 mm in length from FLO15–09/2) were handpicked under a binocular microscope from the sieved and washed <2 mm size fraction. Crystals were loaded into wells in 18 mm-diameter aluminium sample discs for neutron irradiation, along with the Alder Creek sanidine age standard (ACs-2)<sup>43</sup> as the neutron fluence monitor. In this study, we report our age determinations relative to the recently published and astronomically calibrated 1.185 Ma value for ACs-2 (ref. 44). Neutron irradiation was performed in two batches (QL-OSU-39 and QL-OSU-42), each with a duration of 15 min, in the cadmium-shielded CLICIT facility at the Oregon State University TRIGA reactor. Argon isotopic analyses of gas released during the CO<sub>2</sub> laser single crystal fusion experiments (Supplementary Information Table 5) were made on a fully automated Nu Instruments Noblesse multi-collector noble-gas mass spectrometer, using procedures documented previously<sup>14,45</sup>. Reconnaissance isotopic measurements of small gas aliquots released from feldspar crystals during an initial low-temperature heating step, using a low-power defocused laser beam, allowed the identification of scarce, K-rich grains within a population dominated by Ca-rich, K-poor plagioclase; the latter not being amenable to <sup>40</sup>Ar/<sup>39</sup>Ar dating because of small signal size. Fusion experiments on eight of these relatively K-rich grains identified by this prospecting method yielded small but measurable amounts of <sup>40</sup>Ar ( $7.5 \times 10^{-17}$  to  $1.3 \times 10^{-15}$  mol <sup>40</sup>Ar) and <sup>39</sup>Ar. Sample gas cleanup was through an all-metal extraction line, equipped with a –130°C cold trap (to remove H<sub>2</sub>O) and two water-cooled SAES GP-50 getter pumps (to absorb reactive gases). Argon isotopic analyses of unknowns, blanks, and monitor minerals were carried out in identical fashion during a fixed period of 400 s in 14 data acquisition cycles. <sup>40</sup>Ar and <sup>39</sup>Ar were measured on the high-mass ion counter, <sup>38</sup>Ar and <sup>37</sup>Ar on the axial ion counter and <sup>36</sup>Ar on the low-mass ion counter, with baselines measured every third cycle. Measurement of the <sup>40</sup>Ar, <sup>38</sup>Ar, and <sup>36</sup>Ar ion beams was carried out simultaneously, followed by sequential measurement of <sup>39</sup>Ar and <sup>37</sup>Ar. Beam switching was achieved by varying the field of the mass spectrometer magnet and with minor adjustment of the quad lenses. Data acquisition and reduction was performed using the program 'Mass Spec' (A. Deino, Berkeley Geochronology Center). J-values for unknowns were calculated by using a plane-fitting algorithm in the Mass Spec software applied to ACs-2 standard data from symmetrically distributed sample wells across the aluminium irradiation disk (Supplementary Information Table 5). Detector intercalibration and mass fractionation corrections were made using the weighted mean of a time series of measured atmospheric argon aliquots delivered from a calibrated air pipette<sup>14,45</sup>. Sample sets were bracketed by both air pipette and ACs-2 to monitor for possible instrumental drift. Blanks were measured at least once before each run of an unknown and had typical values of  $<1.5 \times 10^{-16}$  mol <sup>40</sup>Ar and  $<1 \times 10^{-18}$  mol <sup>36</sup>Ar for QL-OSU-39 experiments, and  $2\text{--}3 \times 10^{-16}$  mol <sup>40</sup>Ar and  $<2.5 \times 10^{-18}$  mol <sup>36</sup>Ar for those of QL-OSU-42. The higher blanks of the QL-OSU-42 experiments can be attributed to the venting and baking of the Noblesse mass-spectrometer after the completion of QL-OSU-39 analyses. Decay and other constants, including correction factors for interference isotopes produced by nucleogenic reactions, are given in Supplementary Information Table 5.

**Laser ablation U-series analyses.** Laser ablation U-series analyses were carried out on two fragments of the same hominin incisor ((SOA-MM6) sample number 3543A: a fragment of the tooth crown, including portions of dentine and enamel tissues; and sample number 3543B, a cross-section of the root of SOA-MM6 only), as well as on two *Stegodon florensis* molars (sample number 3541, and a molar broken into two fragments, which we subdivided respectively into sample number 3542, and sample number 3544). The experimental setup, measurement conditions, and data evaluation followed principles and procedures described in ref. 46. The dentine and enamel tissues of sample number 3543A were analysed on different tracks, whereas three tracks were analysed across the root section (sample number 3543B). The *Stegodon* molars were analysed along several tracks that cut across the dental tissues. No individual age calculation was carried when the U-concentrations were below about 0.5 ppm and detrital <sup>232</sup>Th was observed (elemental U/Th ratios below 300). The analytical data of the enamel and dentine sections were integrated to provide the data input for the ESR age calculations. All results are shown in Supplementary Information Table 6. ESR dating was also performed on the two *S. florensis* molars noted above. The fossil teeth were prepared following a standard ESR dating procedure for enamel powder<sup>47</sup>. The grey

enamel layer was mechanically separated from the other dental tissues, and both inner and outer surfaces were removed with a dental drill to eliminate the volume that had received an external alpha dose. The samples were then ground and sieved to recover the size fraction of  $<200\mu\text{m}$ . Dose evaluation used the multiple aliquot additive dose method. The powder was split into several aliquots and irradiated up to 4019 Gy with a Gammacell 1000 Cs-137 gamma source. ESR measurements were carried out with a Bruker Elexsys 500 spectrometer, using the following acquisition parameters: 3–5 scans, 2 mW microwave power, 1024 points resolution, 12 mT sweep width, 50 kHz modulation frequency, 0.1 mT modulation amplitude, 20 ms conversion time and 5 ms time constant. The ESR intensities were extracted from T1-B2 peak-to-peak amplitudes of the ESR signal<sup>48</sup>, and then normalized on the number of scans and mass. All aliquots of a given sample were measured within a short time interval ( $<1\text{ h}$ ). This procedure was repeated twice over two successive days without removing the enamel from the ESR tubes between measurements in order to evaluate measurement precision and thus  $D_E$  reproducibility: the latter was found to be excellent, with a variability of  $<3\%$  between the two repeated measurements. Fitting procedures were carried out with the Microcal OriginPro 9.1 software using a Levenberg-Marquardt algorithm by chi-square minimisation. Data were weighted by the inverse of the squared ESR intensity ( $1/I^2$ ) (ref. 49). Final equivalent dose ( $D_E$ ) values were obtained by fitting a single saturating exponential (SSE) through the pooled ESR intensities obtained from the two repeated measurements. Given the magnitude of the  $D_E$  values (between 400 Gy and 600 Gy), a maximum irradiation dose ( $D_{\text{max}}$ ) of 4019 Gy was used, so that  $D_{\text{max}}/D_E$  ratio remains between 5 and 10 as recommended in a recent study<sup>50</sup> to ensure reliable fitting. The final dose response curves (DRCs) are shown in Extended Data Fig. 7f, g. For the dose rate calculations, the following parameters were used: an alpha efficiency of  $0.13 \pm 0.02$  (ref. 51), Monte-Carlo beta attenuation factors from ref. 52, dose-rate conversion factors from ref. 53, an estimated water content of  $5 \pm 3\text{ wt.}\%$  in dentine and  $20 \pm 5\text{ wt.}\%$  in sediment. U and Th and K concentrations in raw sediment were determined by ICP-OES and ICP-MS analysis on samples collected within Layer II (Supplementary Information Table 7). The mean radioelement concentration values were used for the age calculations. Combined U-series/ESR ages were calculated with DATA, a DOS-based programme<sup>54</sup> using the US model defined in ref. 55, and considering the following geometry: sediment/brown enamel/grey enamel/dentine. Further details about this dating method as applied to fossil teeth may be found elsewhere<sup>56</sup>. The results of the age calculations are shown in Supplementary Information Table 8.

31. Westgate, J. A. Isothermal plateau fission-track ages of hydrated glass shards from silicic tephra beds. *Earth Planet. Sci. Lett.* **95**, 226–234 (1989).
32. Westgate, J. A. Fission track dating of volcanic glass, in *Encyclopedia of Scientific Dating Methods* (eds. Rink, W. J. & Thompson, J.) 1–60 (Springer Dordrecht, 2014).
33. Westgate, J. A., Naeser, N. D. & Alloway, B. V. Fission-track dating, in *The Encyclopedia of Quaternary Science* (ed. Elias, S.) 643–662 (Elsevier, 2013).
34. Gansecki, C. A., Mahood, G. A. & McWilliams, M. New ages for the climactic eruptions of Yellowstone: single-crystal  $^{40}\text{Ar}/^{39}\text{Ar}$  dating identifies contamination. *Geology* **26**, 343–346 (1998).
35. Sandhu, A. S. & Westgate, J. A. The correlation between reduction in fission-track diameter and areal track density in volcanic glass shards and its application in dating tephra beds. *Earth Planet. Sci. Lett.* **131**, 289–299 (1995).
36. Laurenzi, M. A. *et al.*  $^{40}\text{Ar}/^{39}\text{Ar}$  laser probe dating of the Central European tektite-producing impact event. *Meteorit. Planet. Sci.* **38**, 887–893 (2003).
37. Laurenzi, M. A. *et al.* New constraints on ages of glasses proposed as reference materials for fission-track dating. *Geostand. Geoanal. Res.* **31**, 105–124 (2007).
38. Pearce, N. J. G. *et al.* Individual glass shard trace element analyses confirm that all known Toba tephra reported from India is from the c. 75-ka Youngest Toba eruption. *Journ. Quat. Sci.* **29**, 729–734 (2014).
39. Westgate, J. A. & Pearce, N. J. G. Quaternary tephrochronology of the Toba tuffs and its significance with respect to archaeological studies in peninsular India, in *Issues in Indian Archaeology: Prehistory to Early History* (ed. Korisettar, R.) (Primus Books New Delhi, in press).
40. Pearce, N. J. G. *et al.* A compilation of new and published major and trace element data for NIST SRM 610 and NIST SRM 612 glass reference materials. *Geost. Newslet.* **21**, 115–144 (1997).
41. Pearce, N. J. G. *et al.* Trace-element analysis by LA-ICP-MS: the quest for comprehensive chemical characterisation of single sub- $10\mu\text{m}$  volcanic glass shards. *Quat. Int.* **246**, 57–81 (2011).
42. Jochum, K. P. *et al.* MPI-DING reference glasses for in-situ microanalysis: new reference values for element concentrations and isotope ratios. *Geochem. Geophys. Geosys.* **7**, Q02008 (2006).
43. Nomande, S. *et al.* Alder Creek sandstone (ACS-2): a Quaternary  $^{40}\text{Ar}/^{39}\text{Ar}$  dating standard tied to the Cobb Mountain geomagnetic event. *Chem. Geol.* **218**, 315–338 (2005).
44. Rivera, T. A., Storey, M., Schmitz, M. D. & Crowley, J. L. Age intercalibration of  $^{40}\text{Ar}/^{39}\text{Ar}$  sanidine and chemically distinct U/Pb zircon populations from the Alder Creek Rhyolite Quaternary geochronology standard. *Chem. Geol.* **345**, 87–98 (2013).
45. Storey, M., Roberts, R. G. & Saidin, M. Astronomically calibrated  $^{40}\text{Ar}/^{39}\text{Ar}$  age for the Toba supereruption and global synchronization of late Quaternary records. *Proc. Natl Acad. Sci. USA* **109**, 18684–18688 (2012).
46. Grün, R. *et al.* Laser ablation U-series analysis of fossil bones and teeth. *Palaeogeogr. Palaeoclimatol.* **416**, 150–167 (2014).
47. Duval, M. *et al.* The challenge of dating Early Pleistocene fossil teeth by the combined uranium series–electron spin resonance method: the Venta Micena palaeontological site (Orce, Spain). *J. Quaternary Sci.* **26**, 603–615 (2011).
48. Grün, R. Methods of dose determination using ESR spectra of tooth enamel. *Radiat. Meas.* **32**, 767–772 (2000).
49. Grün, R. & Brumby, S. The assessment of errors in past radiation doses extrapolated from ESR/TL dose-response data. *Radiat. Meas.* **23**, 307–315 (1994).
50. Duval, M. & Grün, R. Are published ESR dose assessments on fossil tooth enamel reliable? *Quat. Geochronol.* **31**, 19–27 (2016).
51. Grün, R. & Katzenberger-Apel, O. An alpha irradiator for ESR dating. *Ancient TL* **12**, 35–38 (1994).
52. Marsh, R. E. *Beta-gradient Isochrons Using Electron Paramagnetic Resonance: Towards a New Dating Method in Archaeology*. MSc thesis, McMaster University (1999).
53. Adamiec, G. & Aitken, M. J. Dose-rate conversion factors: update. *Ancient TL* **16**, 37–50 (1998).
54. Grün, R. The DATA program for the calculation of ESR age estimates on tooth enamel. *Quat. Geochronol.* **4**, 231–232 (2009).
55. Grün, R., Schwarcz, H. P. & Chadam, J. M. ESR dating of tooth enamel: coupled correction for U-uptake and U-series disequilibrium. *Nucl. Tracks Radiat. Meas.* **14**, 237–241 (1988).
56. Duval, M. Electron Spin Resonance (ESR) dating of fossil tooth enamel, in *Encyclopedia of Scientific Dating Methods* (eds. Rink, W. J. & Thompson, J.) 1–11 (Springer Dordrecht, 2015).
57. Deino, A. & Potts, R. Age-probability spectra for examination of single-crystal  $^{40}\text{Ar}/^{39}\text{Ar}$  dating results: examples from Olorgesailie, southern Kenya Rift. *Quat. Int.* **13–14**, 47–53 (1992).
58. Powell, R., Hergt, J. & Woodhead, J. Improving isochron calculations with robust statistics and the bootstrap. *Chem. Geol.* **185**, 191–204 (2002).
59. Lee, J.-Y. *et al.* A redetermination of the isotopic abundances of atmospheric Ar. *Geochim. Cosmochim. Acta* **70**, 4507–4512 (2006).
60. Moore, M. W. *et al.* Continuities in stone flaking technology at Liang Bua, Flores, Indonesia. *J. Hum. Evol.* **57**, 503–526 (2009).



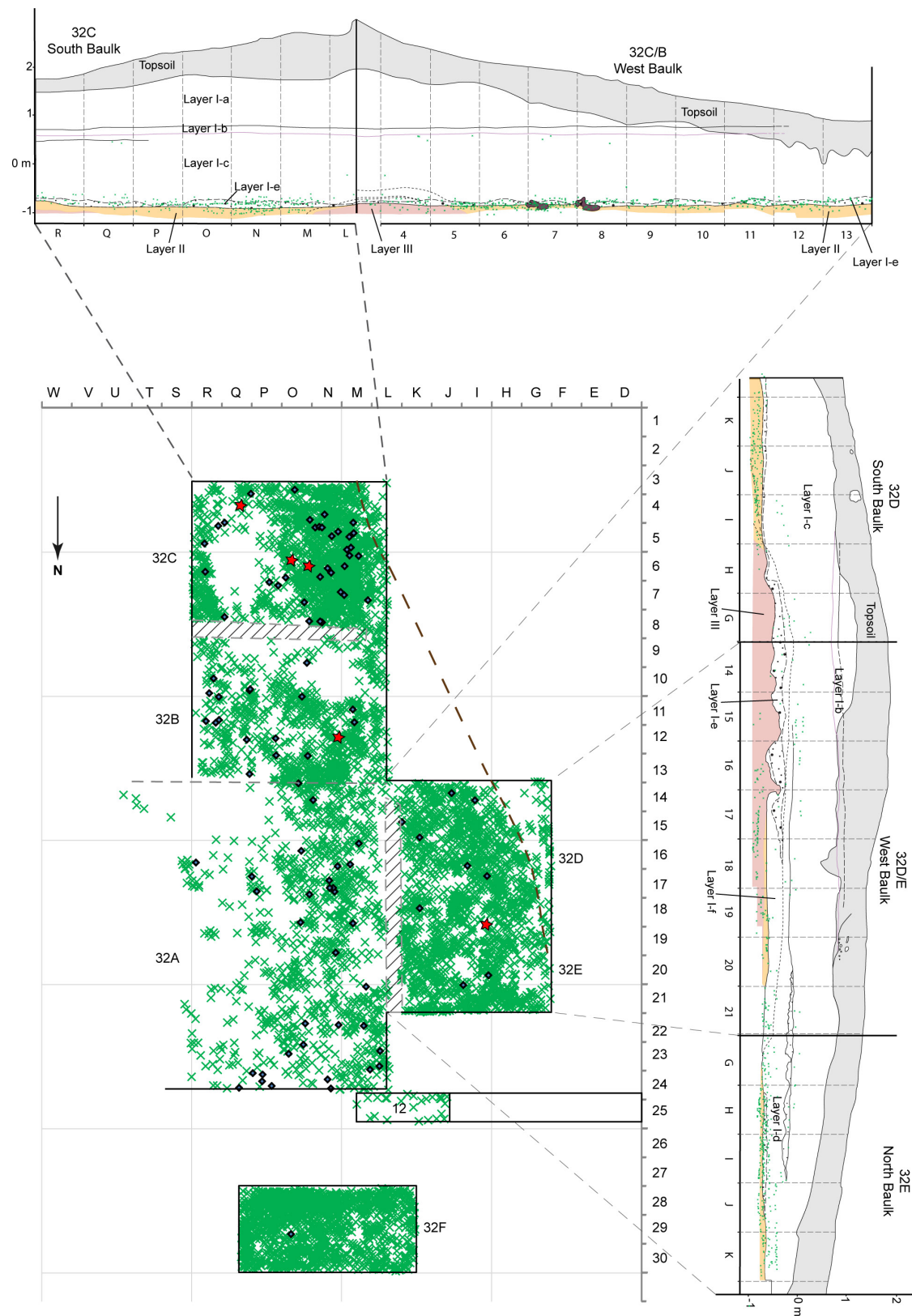


#### Extended Data Figure 1 | Hominin fossil find-locality at Mata Menge.

**a**, View of Excavation 32 (trench E-32) in 2014, taken towards the north-north-west. The dip slope visible in the background is the eastern flank of the Welas Caldera, which was the source for many of the volcanic products deposited in the So'a Basin. **b**, trench E-32A to E viewed towards the southwest, in October 2015. **c**, E-32D to E-32E viewed towards the southwest. The irregular erosional upper surface of the reddish brown palaeosol (Layer III) formed the hardened bedding of a small stream.

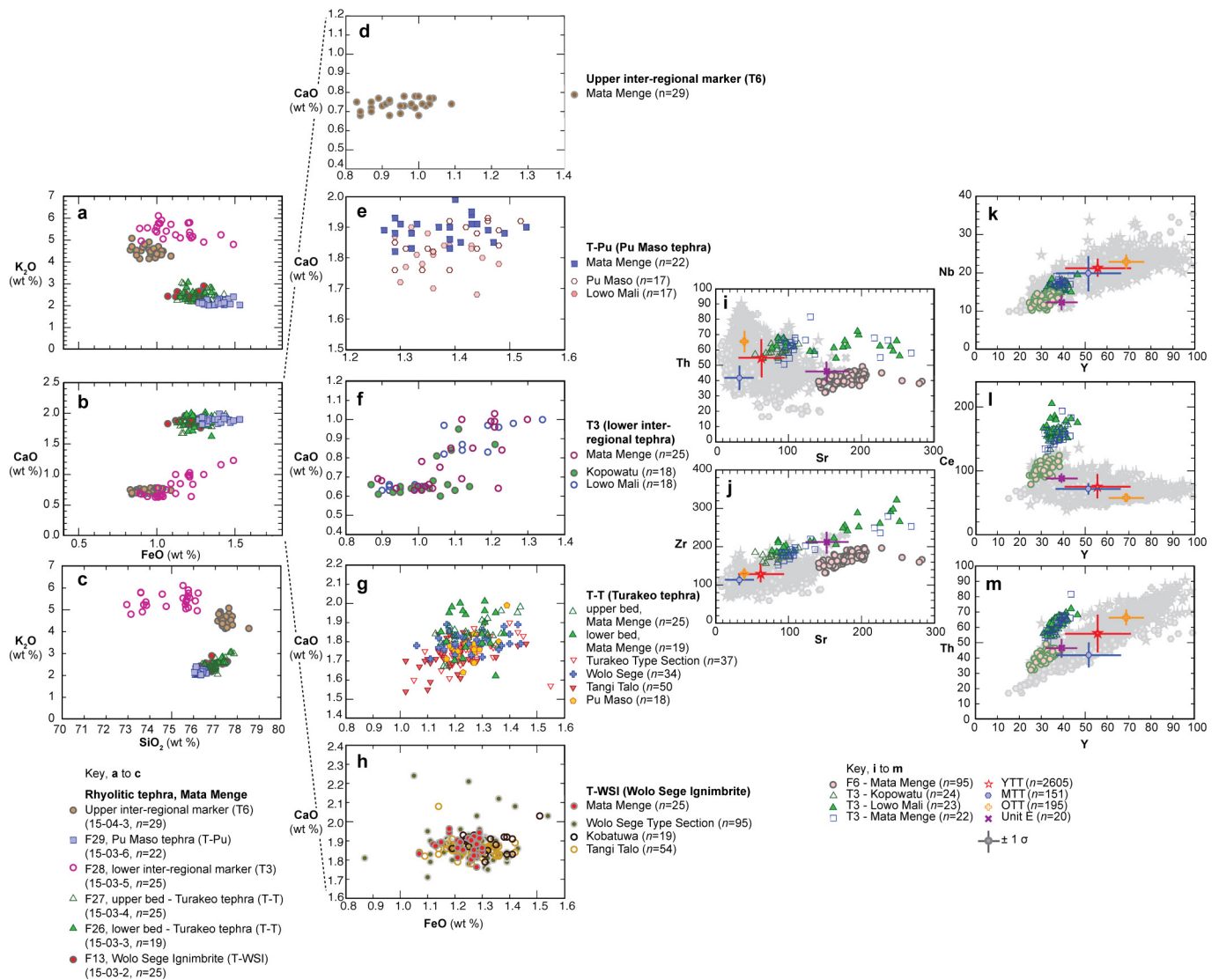
The sandy fossil-bearing Layer II infills depressions formed on this bedding surface. A sequence of mudflows (Layer I/a–f) rapidly covered the entire river bedding and its exposed banks. **d**, Mould of a freshwater gastropod (*Cerithidea*) from sandy Layer II. **e**, Detail of the locally developed, gradual boundary between sandy Layers II and muddy Layer I. Note the abundance of muddy rip clasts around the transition. At other places, the boundary is sharp. **f**, West baulk of E-32C. Large *Stegodon florensis* bones occur at the boundary between Layers II and I.





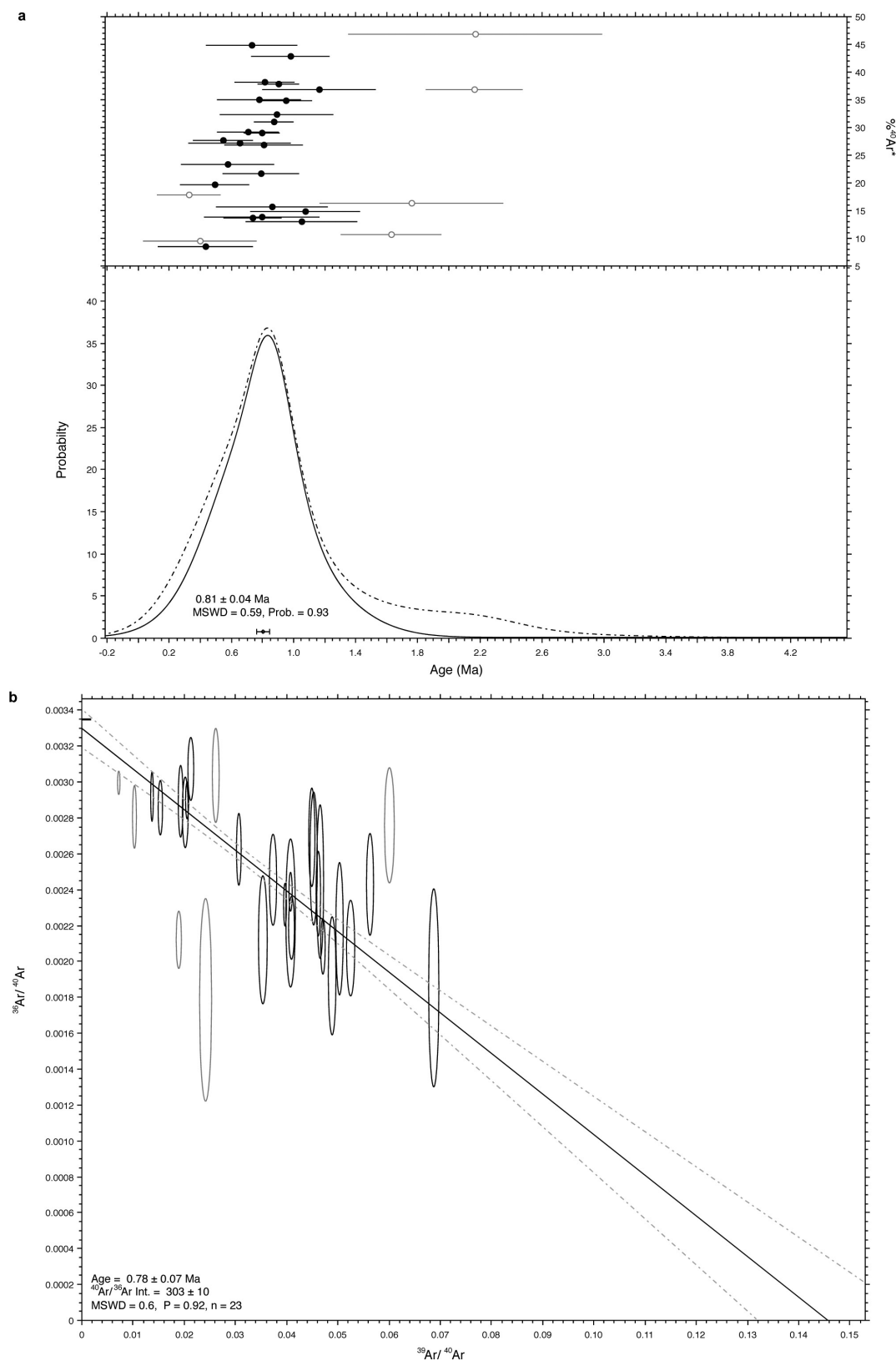
**Extended Data Figure 2 | Plan and baulk profiles of Excavation 32A-F showing distribution of finds.** The horizontal plan (lower left corner) shows the horizontal coordinates of individual fossil finds (green crosses) and stone artefacts (blue diamonds). The position of hominin fossils is indicated with red stars. In the trench baulk profiles (top and right) only

the projected positions of fossil finds occurring within one meter of the baulks are plotted. All hominin fossils were recovered from the top of sandy Layer II. The basal part of the mudflow unit (Layers Ia–e) also contains fossils, stone artefacts, gastropods, and pebbles. The thick brown dotted line indicates the western margin of the ancient stream bed.



**Extended Data Figure 3 | Glass shard geochemistry.** a–c, Selected major element compositions (weight percent FeO vs  $K_2O$ , FeO vs  $CaO$  and  $SiO_2$  vs  $K_2O$ ) of glass shards from key rhyolitic pyroclastic density current (PDC) and airfall deposits at Mata Menge. d–h, Weight percent FeO versus  $CaO$  composition of glass shards from key rhyolitic pyroclastic density current (PDC) and airfall deposits at Mata Menge (in stratigraphic sequence from youngest to oldest) compared with correlatives from adjacent So'a Basin sites. While the major element glass compositions of T-WSI, T-T and T-Pu are all geochemically indistinguishable (they are most likely from the same eruptive source) the major element data for each of the tephra consistently occupies different overlapping fields. Moreover, while subtle geochemical differences exist between T-WSI, T-T, and T-Pu, these tephra can also be readily distinguished in the field by a combination of stratigraphic position and association, as well as by morphological expression. i, j, Selected trace element compositions Sr versus Th and Zr, and k–m, Y versus Nb, Ce and Th of glass shards from T3 correlatives at Mata Menge, Lowo Mali and Kopowatu as well as T6 (uppermost

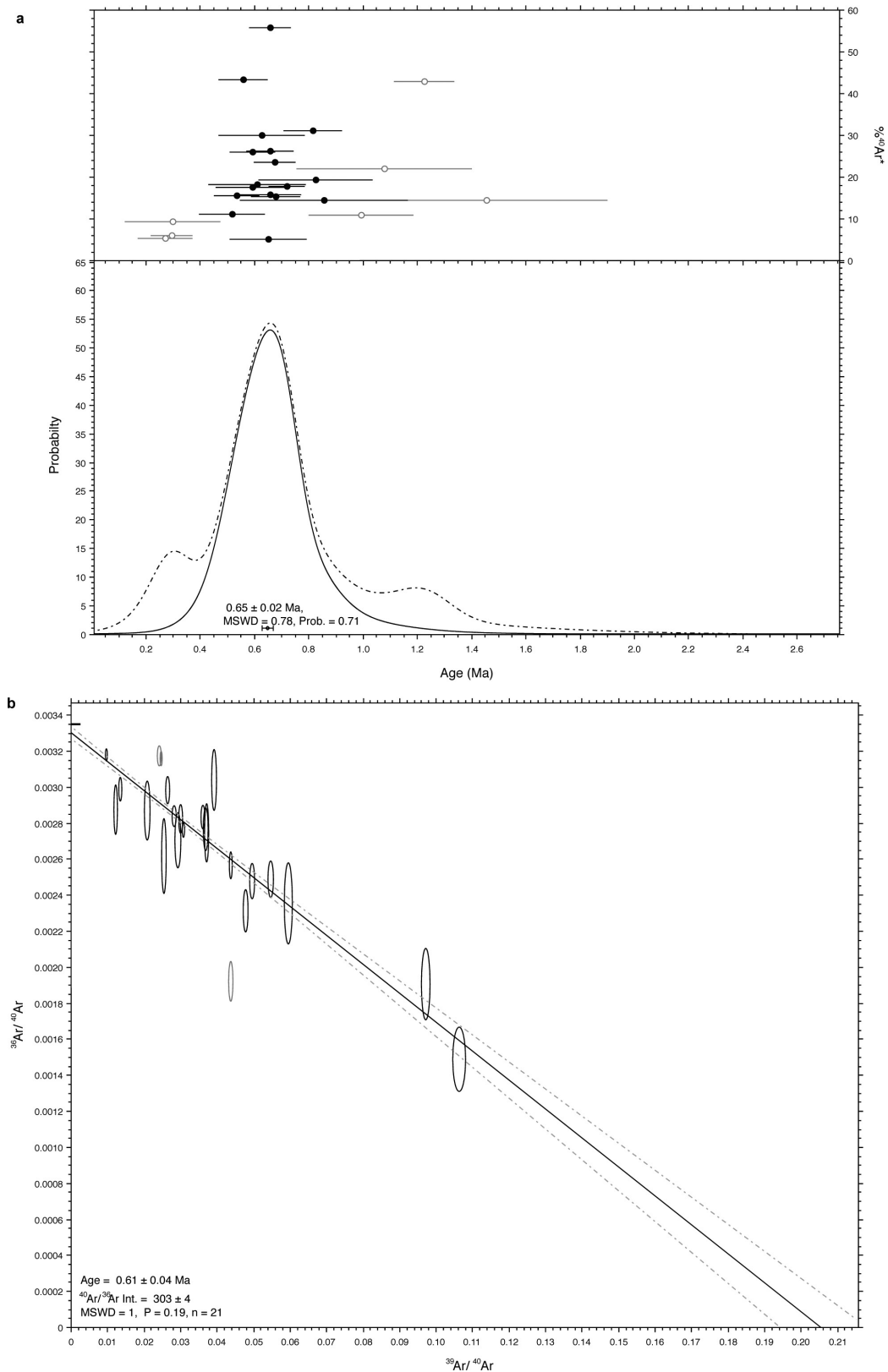
inter-regional marker) from Mata Menge. All trace element concentrations are in ppm unless otherwise stated. The data are plotted against equivalent elemental mean and standard deviation (represented as  $\pm 1\sigma$  error bars) reference data from potential distal tephra correlatives (that is, Youngest Toba Tuff (YTT), Middle Toba Tuff (MTT), Oldest Toba Tuff (OTT) and Unit E from ODP-758) acquired on the same instrument using the same standards and under the same analytical conditions<sup>41,42</sup>. (YTT data source: Pearce, N. J. G., Westgate, J. A. & Gatti, E., Multiple magma batches recorded in tephra deposits from the Toba complex, Sumatra. V51F-3102, AGU Fall Meeting, San Francisco, 14–18 December 2015). Trace element data indicate that the upper (T6) and lower (T3) inter-regional marker beds occurring at Mata Menge cannot be geochemically related to any known Toba-sourced tephra. On this basis, the eruptive sources of T6 and T3 currently remain unknown. However, this absence of eruptive source does not diminish their importance within the overall So'a Basin stratigraphy.



**Extended Data Figure 4 |  $^{40}\text{Ar}/^{39}\text{Ar}$  dating results.** **a**, Age probability plot for single crystal laser fusion data for hornblende from the Pu Maso ignimbrite (sample FLO-15-15; Supplementary Information Table 5); the vertical scale is a relative probability measure of a given age occurring in the sample<sup>57</sup>. We applied an outlier-rejection scheme to the main population to discard ages with normalized median absolute deviations of  $>1.5$  (ref. 58) and these are shown as open circles.  $\%^{40}\text{Ar}^*$  refers to the proportion of radiogenic  $^{40}\text{Ar}$  released for individual analyses. The

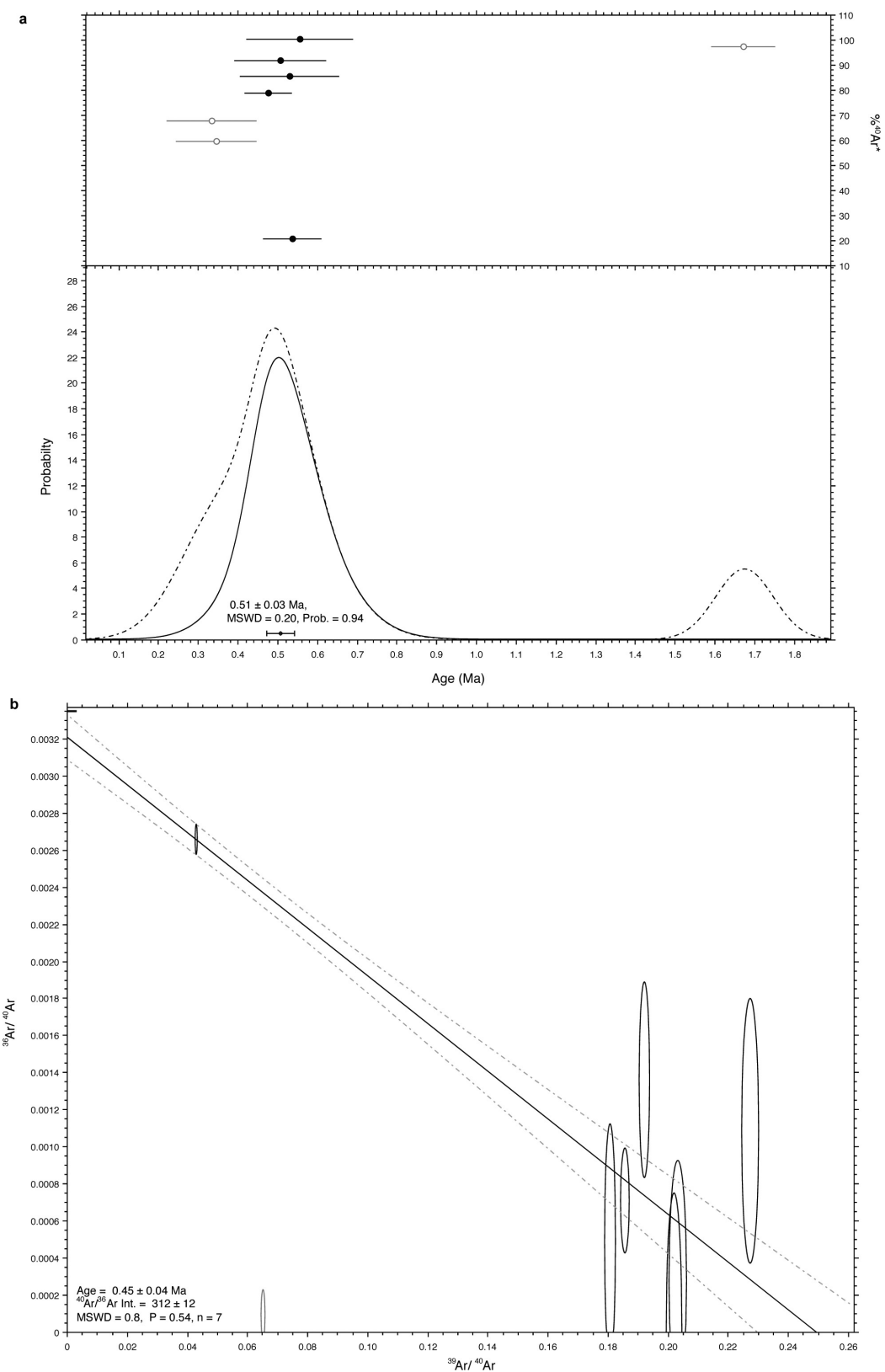
weighted mean age of the filtered hornblende data for the Pu Maso ignimbrite is  $0.81 \pm 0.04$  Ma ( $1\sigma$ ; mean square of the weighted deviates (mswd) = 0.59, prob = 0.93;  $n = 23/29$ ). **b**, An inverse isochron plot for these 23 analyses gives a statistically overlapping age of  $0.78 \pm 0.07$  Ma ( $1\sigma$ ; mswd = 0.6, prob. = 0.92). The  $^{40}\text{Ar}/^{36}\text{Ar}$  intercept of  $303 \pm 10$  is statistically indistinguishable from the atmospheric ratio of  $298.6 \pm 0.3$  (ref. 59), thus supporting the more precise weighted mean age result.





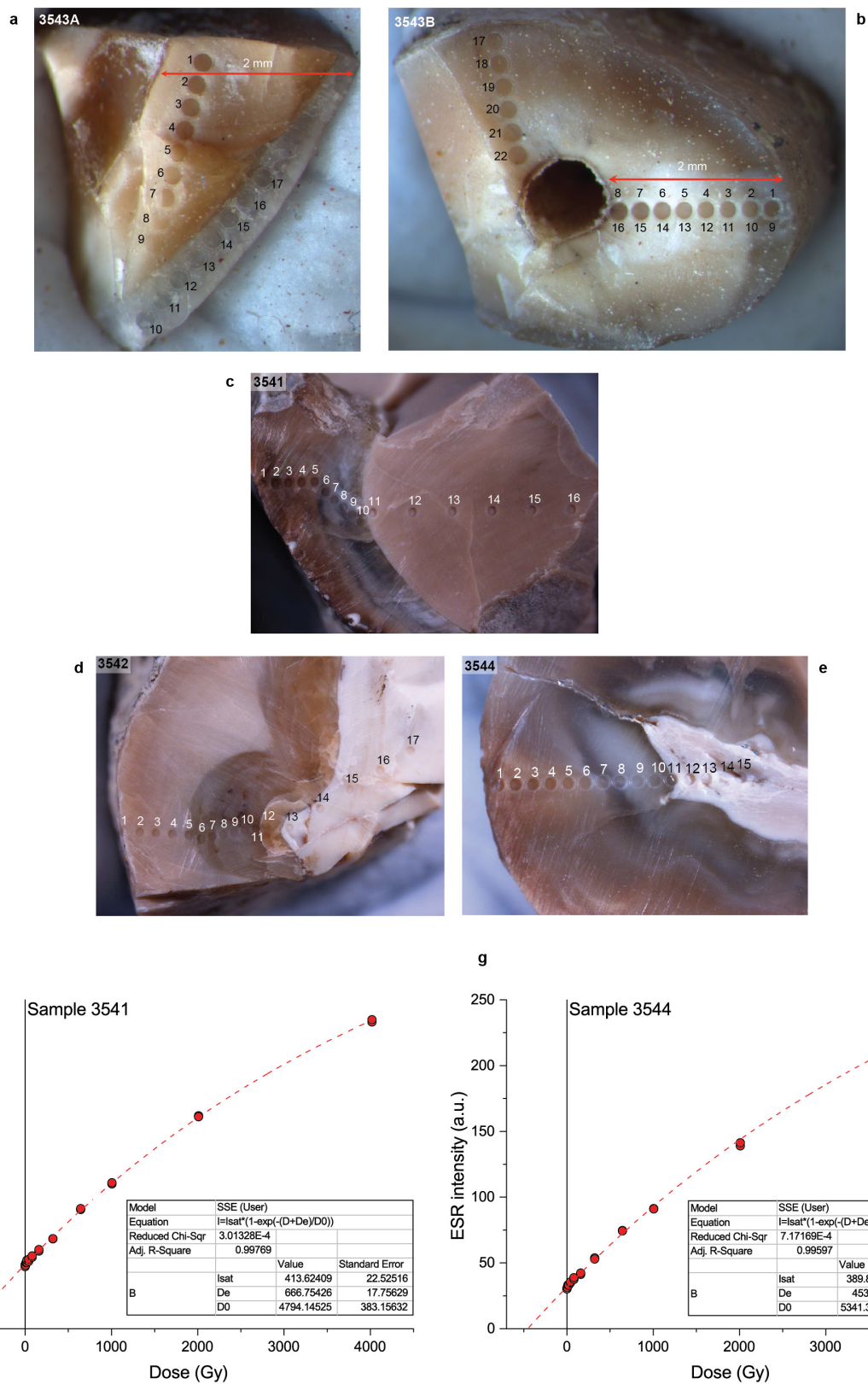
**Extended Data Figure 5** |  $^{40}\text{Ar}/^{39}\text{Ar}$  dating results. **a**, Age probability plot for single crystal laser fusion data for hornblende from the PGT-2 tephra (sample T XII 252-261; Supplementary Information Table 5).  $^{40}\text{Ar}^*$  ranges from <10% to nearly 60%. The weighted mean age of the

filtered hornblende data for the PGT-2 tephra is  $0.65 \pm 0.02$  Ma (1 $\sigma$ ; mean square of the weighted deviates (mswd) = 0.78, prob = 0.71;  $n = 17/24$ ). **b**, An inverse isochron plot gives a statistically overlapping, but less precise, age of  $0.61 \pm 0.04$  Ma (1 $\sigma$ ; mswd = 1,  $P = 0.19$ ).



**Extended Data Figure 6 |  $^{40}\text{Ar}/^{39}\text{Ar}$  dating results.** **a**, Age probability plot for single crystal laser fusion data for anorthoclase from the T6 upper inter-regional rhyolitic tephra (sample FLO15-09/2; Supplementary Information Table 5).  $^{40}\text{Ar}^*$  ranges from 20% to nearly 100%. The

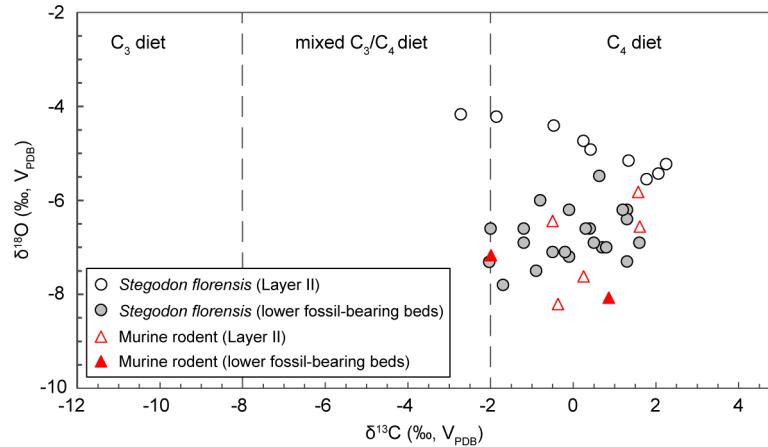
weighted mean age of the filtered feldspar data for the T6 tephra is  $0.51 \pm 0.03$  Ma ( $1\sigma$ ;  $\text{mswd} = 0.20$ ,  $\text{prob} = 0.94$ ;  $n = 5/8$ ). An inverse isochron plot (**b**) gives a statistically overlapping, but less precise, age of  $0.45 \pm 0.04$  Ma ( $1\sigma$ ;  $\text{mswd} = 0.8$ ,  $P = 0.54$ ).



**Extended Data Figure 7 | U-series and ESR samples and dating results.**  
**a, b**, Hominin incisor (SOA-MM6) crown and root samples (number 3543A and number 3543B, respectively) from Layer II, Mata Menge.  
**c–e**, U-series laser tracks for *Stegodon* molar samples from Layer II.  
**f, g**, Dose response curves obtained for the two powder enamel samples

from number 3541 and number 3544, respectively. Fitting was carried out with a SSE function through the pooled mean ESR intensities derived from each repeated measurement. Given the magnitude of the  $D_E$  values, the correct  $D_E$  value was obtained for  $5 > D_{max}/D_E > 10$  (ref. 50).



**a****b**

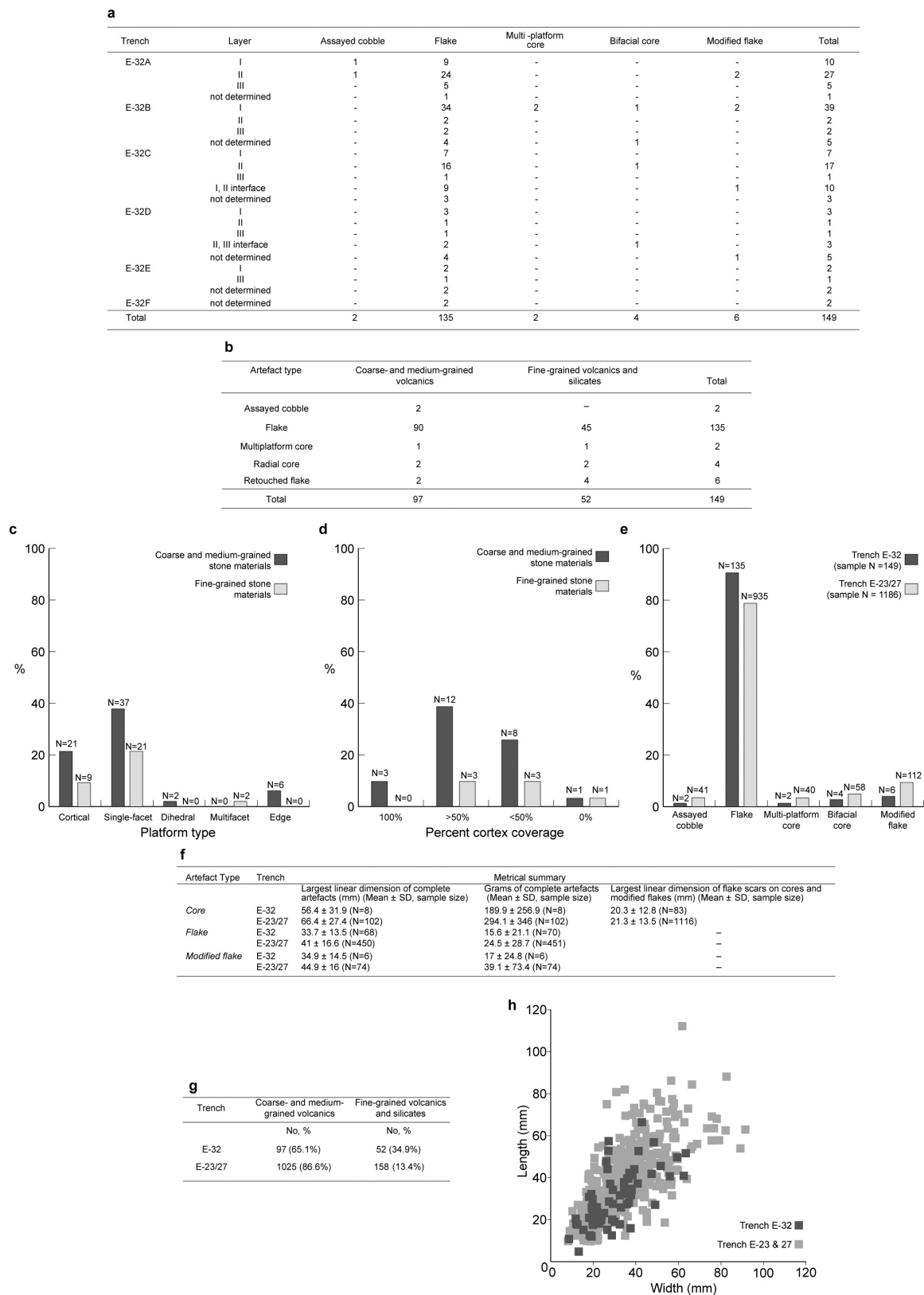
	<i>Stegodon</i> , lower levels	Murine rodent, lower levels	<i>Stegodon</i> , upper level (Layer II)	Murine rodent, upper level (Layer II)
<i>Stegodon</i> , lower levels (n=22)		0.251	0.281	0.3792
Murine rodent, lower levels (n=2)	0.251		0.2374	0.3329
<i>Stegodon</i> , upper level (Layer II; n=9)	0.281	0.2374		0.8542
Murine rodent, upper level (Layer II; n=5)	0.3792	0.3329	0.8542	

**c**

	<i>Stegodon</i> , lower levels	Murine rodent, lower levels	<i>Stegodon</i> , upper level (Layer II)	Murine rodent, upper level (Layer II)
<i>Stegodon</i> , lower levels (n=22)		1.00	<b>0.00287</b>	1
Murine rodent, lower levels (n=2)	1.00		0.6427	1
<i>Stegodon</i> , upper level (Layer II; n=9)	<b>0.00287</b>	0.6427		0.07229
Murine rodent, upper level (Layer II; n=5)	1	1	0.07229	

**Extended Data Figure 8 | Carbon and oxygen isotope analysis of dental enamel.** **a**,  $\delta^{13}\text{C}$  and  $\delta^{18}\text{O}$  values of *Stegodon florensis* and murine rodent tooth enamel. All but one of the  $\delta^{13}\text{C}$  ratios correspond to a  $\text{C}_4$  diet, indicating that the analysed *Stegodon* and murine rodents were predominantly grazers. The positive shift observed in  $\delta^{18}\text{O}$  of the younger *Stegodon* samples (from the hominin-bearing Layer II) is more difficult to interpret with the limited data available, but could mean a distinct

source of drinking water (i.e., run-off versus lacustrine) and/or warmer conditions. **b**, Benferroni corrected  $P$  values for a pairwise Mann–Whitney statistical analysis to test for similarity of  $\delta^{13}\text{C}$  between subsamples. **c**, Benferroni corrected  $P$  values for a pairwise Mann–Whitney statistical analysis to test for similarity of  $\delta^{18}\text{O}$  between subsamples;  $P$  values showing significant differences in median values are in bold.



**Extended Data Figure 9 | Analytical data for the Mata Menge stone technology.** **a**, Artefact counts and provenance, trench E-32 (artefact definitions after ref. 60). **b**, raw materials used to manufacture the stone tool assemblage, trench E-32. **c**, Platform types on flakes and modified flakes, E-32. Cortical: the blow was struck onto the cortical surface of a cobble. Single-facet: the blow was struck on a scar produced by previous reduction. Dihedral: the blow was struck on the ridge between two scars produced by previous reduction. Multifacet: the blow was struck on the surface of multiple small scars produced by previous reduction. Edge: the blow was struck on the edge of the core and a platform surface is not

retained on the flake. **d**, Cortex coverage on the dorsal surface of complete unmodified flakes, E-32. Percent cortex coverage refers to the proportion of the dorsal surface covered in cortex. **e**, Artefact counts, trenches E-32 and E-23/27 (artefact definitions after ref. 60). **f**, Sizes of artefacts and attributes, E-32 and E-23/27. **g**, Raw materials used to manufacture the stone tool assemblage, E-32 and E-23/27. **h**, Scatterplot of complete flake sizes, E-32 (total sample size  $n = 68$  complete flakes) and E-23/27 ( $n = 443$ ). With regards to raw materials, coarse- and medium-grained materials include andesite, basalt, rhyolite, and tuff. Fine-grained materials include silicified tuff, chalcedony, and opal.

# Environmental Breviatea harbour mutualistic *Arcobacter* epibionts

Emmo Hamann<sup>1,2</sup>, Harald Gruber-Vodicka<sup>3</sup>, Manuel Kleiner<sup>2</sup>, Halina E. Tegetmeyer<sup>1,4</sup>, Dietmar Riedel<sup>5</sup>, Sten Littmann<sup>6</sup>, Jianwei Chen<sup>1,2</sup>, Jana Milucka<sup>6</sup>, Bernhard Viehweger<sup>7</sup>, Kevin W. Becker<sup>7</sup>, Xiaoli Dong<sup>2</sup>, Courtney W. Stairs<sup>8</sup>, Kai-Uwe Hinrichs<sup>7</sup>, Matthew W. Brown<sup>9</sup>, Andrew J. Roger<sup>8</sup> & Marc Strous<sup>1,2,4</sup>

**Breviatea form a lineage of free living, unicellular protists, distantly related to animals and fungi<sup>1,2</sup>. This lineage emerged almost one billion years ago, when the oceanic oxygen content was low, and extant Breviatea have evolved or retained an anaerobic lifestyle<sup>3,4</sup>. Here we report the cultivation of *Lenisia limosa*, gen. et sp. nov., a newly discovered breviate colonized by relatives of animal-associated *Arcobacter*. Physiological experiments show that the association of *L. limosa* with *Arcobacter* is driven by the transfer of hydrogen and is mutualistic, providing benefits to both partners. With whole-genome sequencing and differential proteomics, we show that an experimentally observed fitness gain of *L. limosa* could be explained by the activity of a so far unknown type of NAD(P)H-accepting hydrogenase, which is expressed in the presence, but not in the absence, of *Arcobacter*. Differential proteomics further reveal that the presence of *Lenisia* stimulates expression of known 'virulence' factors by *Arcobacter*. These proteins typically enable colonization of animal cells during infection<sup>5</sup>, but may in the present case act for mutual benefit. Finally, re-investigation of two currently available transcriptomic data sets of other Breviatea<sup>4</sup> reveals the presence and activity of related hydrogen-consuming *Arcobacter*, indicating that mutualistic interaction between these two groups of microbes might be pervasive. Our results support the notion that molecular mechanisms involved in virulence can also support mutualism<sup>6</sup>, as shown here for *Arcobacter* and Breviatea.**

As a cause of genomic innovations and a catalyst of diversification, close interactions between eukaryotes and prokaryotes are driving forces of evolution<sup>7</sup>. The importance of eukaryote–prokaryote interactions is clearly manifested in the remarkable diversity and abundance of bacteria that live in symbiosis with large multicellular eukaryotes such as animals. The basic adaptive requirements for symbiotic bacteria are the capability to recognize and colonize host tissue, evasion of defence mechanisms, replication and ultimately the transfer to new hosts. Bacteria may have first evolved the capability for symbiotic interactions with eukaryotes through associations with ancestral protists<sup>7</sup>. Today, several protist lineages are known that have preserved ancestral eukaryotic features<sup>1,8,9</sup>. Characterizing these lineages and their molecular interactions with bacteria is vital to formulate evidence-based hypotheses for the origin and functions of ancestral bacterial–eukaryote symbioses.

By providing nitrous oxide as the only electron acceptor, and the bacterium *Alteromonas macleodii* as prey bacteria, we enriched an amoeboid flagellate colonized by spiral-shaped bacteria from an anoxic marine tidal-flat sediment. DNA was extracted from the enrichment culture and used for metagenomic sequencing (see below). Phylogenetic analysis of a concatenated sequence alignment

comprising 16 universal eukaryotic genes identified the flagellate as a new species within Breviatea (Fig. 1a), with *Pygsuma biforma* as its closest relative. We designated the protist as the novel genus and species *Lenisia limosa* (see Supplementary Notes for diagnosis, habitat description and etymology).

*Lenisia limosa* gen. et sp. nov.

*L. limosa* is a small, marine amoeboid flagellate with a predatory lifestyle (Fig. 1c–i). Its morphology has both swimming and adherent gliding forms. Adherent cells are 4–9 µm long and 3–4 µm wide. Swimming cells are 4–6 µm long and 3–4 µm wide. For swimming, *L. limosa* beats its two flagella (anterior flagellum 3–8 µm long and posterior flagellum usually 7–19 µm long, two to three times longer), resulting in slow, wobbling locomotion. When it encounters a substrate, it attaches to it, wraps one flagellum around its lateral side, elongates in shape and starts gliding. While gliding, the second flagellum remains detached and assists in the acquisition of prey bacteria. These are captured with small, filamentous pseudopodia (4–12 µm) originating from the ventral side of the cell. Ultrastructural analysis showed the presence of several key features previously identified in other breviate (Extended Data Fig. 1). These features include a complex internal membrane system, centrioles, two basal bodies, a Golgi apparatus, digestive vacuoles as well as mitochondria-related organelles ('hydrogenosomes').

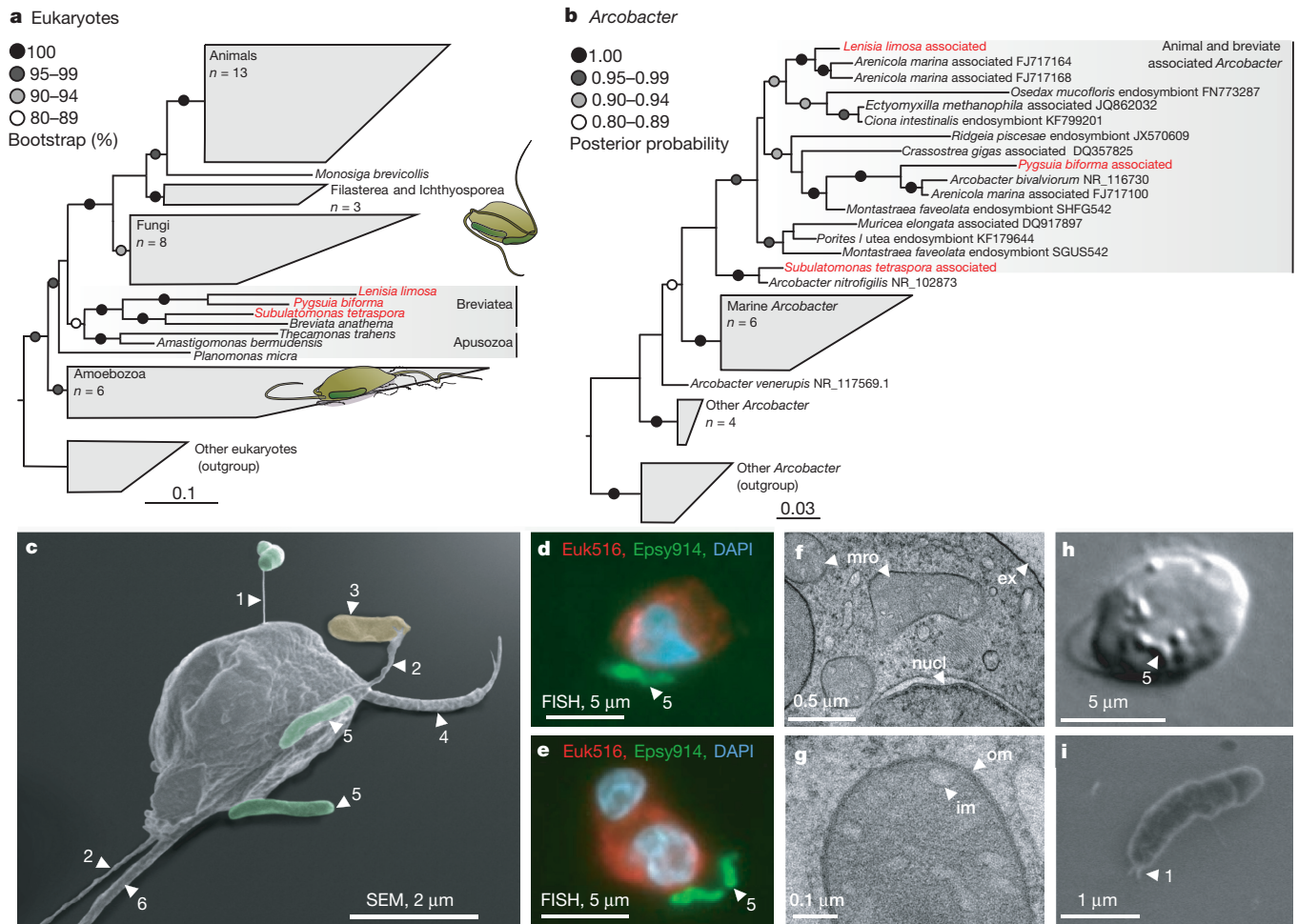
Phylogenetic analysis of bacterial 16S rRNA gene sequences, combined with catalysed reporter deposition fluorescence *in situ* hybridization (CARD-FISH), identified the epibionts of *L. limosa* (typically one to three epibionts per cell) as a species of so far uncultivated Epsilonproteobacteria of the genus *Arcobacter* (Fig. 1b, d, e). *Arcobacter* sp. was closely related to several uncultivated species that were found in association with marine animals. CARD-FISH showed that *Arcobacter* bacteria successfully evaded ingestion, as they were never detected inside cells of *L. limosa*. Other denitrifying bacteria (related to *Colwellia psychrerythraea*) were also detected, both by metagenomic sequencing and by CARD-FISH microscopy, but these did not colonize cells of *L. limosa*.

We found that the symbiosis was facultative for both partners. Provided with dissolved organic carbon, hydrogen and nitrous oxide as the only electron acceptor, *Arcobacter* sp. grew independent of its host. However, in the absence of *L. limosa*, *C. psychrerythraea* became much more abundant than *Arcobacter* (Extended Data Fig. 2). Further evidence for the facultative nature of the symbiosis was obtained by cultivating *L. limosa* without nitrous oxide, which resulted in the loss of its *Arcobacter* symbionts (Extended Data Fig. 2).

Why did *Arcobacter* sp. colonize *L. limosa*? To address this question, we investigated potential metabolic interactions between both partners with combined metagenomics approaches. We performed

<sup>1</sup>Microbial Fitness Group, Max Planck Institute for Marine Microbiology, Celsiusstraße 1, 28359 Bremen, Germany. <sup>2</sup>Department of Geoscience, University of Calgary, Calgary, 2500 University Drive Northwest, Alberta T2N 1N4, Canada. <sup>3</sup>Symbiosis Department, Max Planck Institute for Marine Microbiology, Celsiusstraße 1, 28359 Bremen, Germany. <sup>4</sup>Institute for Genome Research and Systems Biology, Center for Biotechnology, University of Bielefeld, Universitätsstraße 25, 3615 Bielefeld, Germany. <sup>5</sup>Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany. <sup>6</sup>Biogeochemistry Department, Max Planck Institute for Marine Microbiology, Celsiusstraße 1, 28359 Bremen, Germany. <sup>7</sup>MARUM Centre for Marine Environmental Sciences, Bibliothekstraße 1, University of Bremen, 28359 Bremen, Germany. <sup>8</sup>Centre for Comparative Genomics and Evolutionary Bioinformatics, Department of Biochemistry and Molecular Biology, 6299 South Street, Dalhousie University, Halifax, Nova Scotia B3H 4R2, Canada. <sup>9</sup>Department of Biological Sciences, Mississippi State University, Mississippi State, Mississippi 39762, USA.





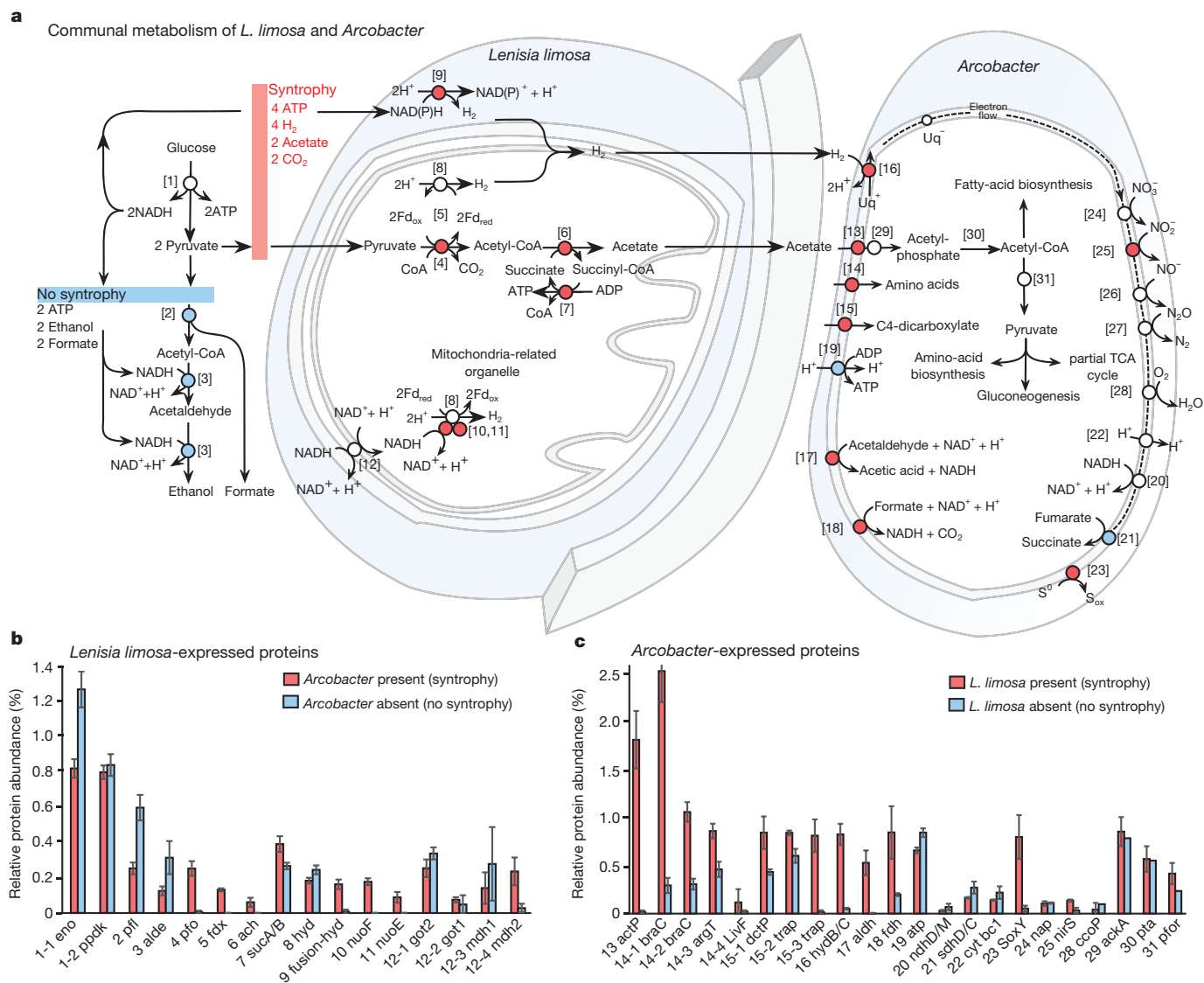
**Figure 1 | Relatives of animal-associated *Arcobacter* colonize Breviatea.** **a**, Maximum likelihood tree of breviatea found in association with *Arcobacter* (in red). **b**, MrBayes tree of *Arcobacter* found in association with animals or breviatea (in red). Dots indicate bootstrap support and posterior probabilities values, respectively. The scale bars represent substitution rate per site. **c**, Scanning electron micrograph of *L. limosa* and associated bacteria. Pilus (1), pseudopodial extensions (2), prey bacteria (3), short anterior flagellum (4), *Arcobacter* (5), long posterior flagellum (6). The background of this image was removed and the bacteria were manually coloured. See Extended Data Fig. 1 for an unmodified

whole-genome and -transcriptome sequencing for the consortium as a whole and performed proteomics for the consortium and for both partners grown separately. Raw sequencing reads were assembled and binned, resulting in a 47 Mb (16 $\times$  coverage) provisional whole-genome sequence for *L. limosa* and a 3 Mb genome (15 $\times$  coverage) for its *Arcobacter* epibiont. For *L. limosa* we performed evidence-driven gene prediction and repeat identification using the assembled contigs and a polyA-tail enriched transcriptome consisting of 48,530 assembled transcripts (22.5 Mb). We predicted 8,146 protein-coding genes covering 15.6% of the host genome. Both genomes were inferred to be nearly complete, with 95% of conserved eukaryotic and 99% of conserved Epsilonproteobacteria genes present (Extended Data Fig. 3).

*L. limosa* encoded a mosaic of genes enabling fermentative ATP production (Fig. 2a). Among genes typically supporting aerobic growth, we identified genes for a partial tricarboxylic acid cycle, an alternative oxidase, as well as a malate-aspartate shuttle. Despite the presence of these genes, *L. limosa* appeared to have lost the capability of oxidative phosphorylation because it lacked Complex IV and the capacity for the biosynthesis of ubiquinol and cytochrome *c*. Further, all subunits for the F-type ATP synthetase (complex V) were absent. Thus, we infer that

the energy metabolism of *L. limosa* depends strictly on fermentative ATP production. We identified two potential pathways for pyruvate oxidation in the genome of *L. limosa*, and proteomics revealed that the expression of either pathway was dependent on the presence/absence of *Arcobacter* sp. In the absence of *Arcobacter*, *L. limosa* appeared to metabolize pyruvate mainly through the activity of pyruvate formate lyase and ethanol dehydrogenase (Fig. 2, gene numbers [2] and [3]). Inside the mitochondria-related organelles, reduced ferredoxin was oxidized by a ferredoxin-dependent hydrogen-evolving hydrogenase (Fig. 2, number [8]). This pathway does enable recycling of cytosolic NADH but is not coupled to the production of additional ATP.

In the presence of *Arcobacter*, *L. limosa* was inferred to switch to a bioenergetically much more efficient metabolism that produces hydrogen not only from ferredoxin but also from NADH. Two distinct enzymes for NADH oxidation were identified and both were only expressed in the presence of the symbiont: a fusion enzyme with an NADH/NADPH-accepting domain, homologous to P450 reductase; and a Fe-hydrogenase domain, which was inferred to produce hydrogen



**Figure 2 | Symbiotic metabolism of *L. limosa* and *Arcobacter*.**

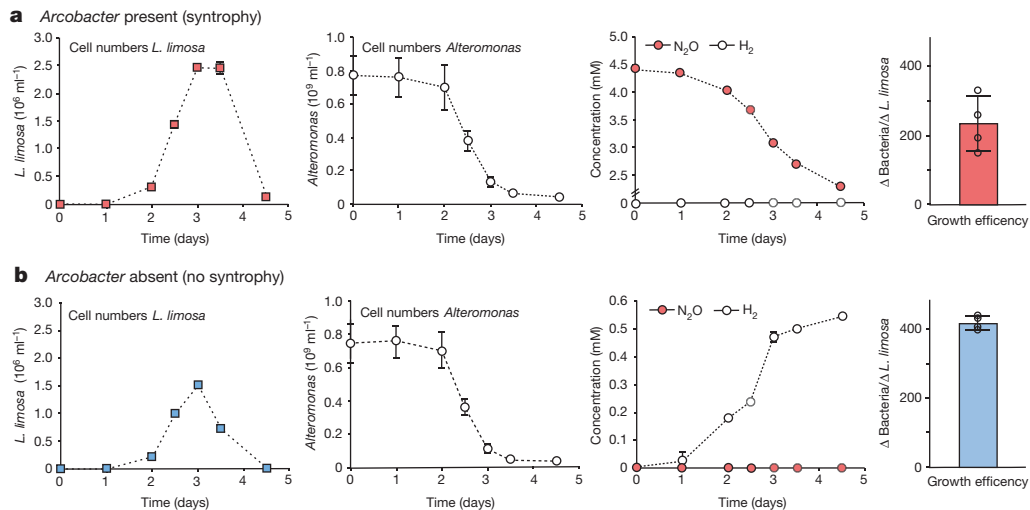
**a**, Symbiotic metabolism of *L. limosa* and *Arcobacter* sp. as inferred from genomics, transcriptomics and proteomics. We identified two main fermentation pathways in *L. limosa*, of which one was coupled to the activity of a NAD(P)H-dependent Fe-hydrogenase (gene number [9], see also Extended Data Fig. 4). The latter pathway theoretically yields two times more ATP and was only expressed in presence of hydrogen oxidizing *Arcobacter*. Numbers correspond to gene names and expression values listed in **b** and **c**. Red circles indicate genes that are more highly expressed under syntrophic conditions; blue circles indicate genes that are more

highly expressed in non-syntrophic conditions. **b**, Expression levels of proteins involved in energy conservation in *L. limosa* in the presence (red) and absence (blue) of *Arcobacter*. **c**, Expression levels of proteins involved in energy conservation and organic carbon uptake expressed by *Arcobacter* in the presence (red) and absence (blue) of *L. limosa*. Error bars, s.d. from three independent experiments (see also Extended Data Fig. 2). If a protein consisted of more than one subunit, the average for all subunits is shown. Subcellular localization of proteins was inferred from the presence of amino (N)-terminal targeting signals. See Supplementary Table 1 for gene accession numbers.

from NADH in the cytosol (Fig. 2, gene number [9], and Extended Data Fig. 4). An enzyme with the same inferred domain structure is present in the breviate *P. biforma*<sup>10</sup>. Inside the mitochondria-related organelles, a NADH dehydrogenase (Fig. 2, gene numbers [10] and [11]) might act together with the ferredoxin-dependent hydrogenase (gene number [8]) to produce hydrogen by electron confurcation<sup>10–12</sup> (Fig. 2). In combination with pyruvate-ferredoxin-oxidoreductase, acetate:succinate CoA-transferase and succinyl-CoA synthetase, NADH-dependent hydrogen production would enable the oxidation of pyruvate to acetate and CO<sub>2</sub> with the production of additional ATP. Expression of this entire metabolism was strongly stimulated in the presence of *Arcobacter* (Fig. 2).

While the production of hydrogen is thermodynamically not prone to product inhibition with electrons from ferredoxin, the production of molecular hydrogen from NAD(P)H only proceeds at low hydrogen

partial pressure (<5.4 μM at NADH/NAD<sup>+</sup> = 10)<sup>13</sup>. The activity of the pathways inferred above therefore requires an active hydrogen sink<sup>13</sup>. Our proteomic analysis suggested that *Arcobacter* acted as a hydrogen sink by expressing a high-affinity, hydrogen-oxidizing Ni/Fe-hydrogenase (Fig. 2, number [16], and Extended Data Fig. 5). At a hydrogen turnover rate of  $9 \times 10^{-18}$  mol H<sub>2</sub> per second per cell of *L. limosa*, two *Arcobacter* epibionts should be able to maintain the hydrogen concentration in the *L. limosa* cytosol at ~5 μM. We also detected the expression of proteins that potentially catalyse the uptake and utilization of all other fermentation products inferred to be produced by *L. limosa*. This includes the anabolic uptake of acetate, as well as the catabolic oxidation of formate. The absence of essential genes supporting autotrophic growth (for example, absence of succinyl-coenzyme-A synthetase and citrate lyase) indicates a general dependency of *Arcobacter* on organic substrates provided by its host.



**Figure 3 | The fitness of *L. limosa* depends on its symbiont.** Syntrophy was enabled by the presence of nitrous oxide acting as electron acceptor for bacterial hydrogen oxidation. **a**, Growth of *L. limosa* in the presence of nitrous oxide (syntrophy) compared with prey abundance (*Alteromonas*) and nitrous oxide concentration. **b**, Growth of *L. limosa* in the absence of nitrous oxide (no syntrophy) compared with prey abundance and hydrogen concentration. Cell numbers of *L. limosa* and concentrations were averaged from two independent experiments per treatment, with the

error bars showing the full range between both measurements. For cell numbers of *Alteromonas*, determined with CARD-FISH, the error bars depict the s.d. of the bacterial cell counts. Growth efficiency was calculated from the difference in bacterial and *L. limosa* cell numbers between two different time points twice during exponential growth. All four individual results are shown as circles; bar heights indicate averages; error bars, s.d. Error bars smaller than data points are not shown. See also Extended Data Fig. 6.

This is consistent with high expression levels observed for putative amino-acid- and fatty-acid-metabolizing enzymes. Our genomic analysis also indicated that hydrogen oxidation could in theory proceed in the presence of electron acceptors other than nitrous oxide. We found all key genes for denitrification, ammonification, as well as the respiration of fumarate and oxygen in the *Arcobacter* genome. Stimulation of growth of *L. limosa*/*Arcobacter* consortia was confirmed experimentally for both oxygen and nitrate (Extended Data Fig. 6g).

Cultivation of *L. limosa* in the absence and presence of its symbionts enabled us to directly measure the fitness effects resulting from the symbiosis. As a fitness indicator, we measured growth rates and growth yields of *L. limosa* in the absence and in presence of *Arcobacter* (Fig. 3). As expected, absence of a syntrophic partner resulted in an accumulation of hydrogen and in a significantly impaired fitness of *L. limosa*. This was apparent from reduced cell numbers, growth rates and growth yields. The growth yield of *L. limosa* was about two times higher during syntrophic growth. Negative effects on the fitness of *L. limosa* were also observed when hydrogen was added directly to the culture or when hydrogen oxidation was abolished through a respiration inhibitor (Extended Data Fig. 6). In combination, our physiological experiments demonstrated true metabolic advantage of *L. limosa*, provided by a hydrogen-scavenging partner, as inferred from combined genomics, transcriptomics and proteomics (Fig. 2).

To understand how *Arcobacter* sp. managed to colonize cells of *L. limosa*, we inspected the genome of the epibiont for the presence of specific genes previously associated with host recognition and attachment. We found that the *Arcobacter* epibiont encoded an almost identical set of 'virulence' genes previously identified in *Arcobacter* and *Campylobacter* pathogens of animals<sup>5</sup>. Among these gene products, the adhesion protein MOMP (*Campylobacter jejuni* major outer membrane protein), CadF and flagellin mediate binding to fibronectin type III, which is found in the extracellular matrix of animal cells<sup>14,15</sup>. The *L. limosa* epibiont was found to express these 'virulence' genes and, for MOMP, flagellin and a subset of chemotaxis proteins, expression was stimulated in the presence of *L. limosa* (Extended Data Fig. 7). Chemotaxis proteins were previously shown to enable pathogenic *Campylobacter* to move into the immediate proximity of their target<sup>16</sup>. Interestingly, we found that *L. limosa* expressed two fibronectin type III domain-containing proteins, potential targets of *Arcobacter*

MOMP, CadF and flagellin (Extended Data Fig. 8). Together, these findings suggest that the mechanism for colonization of *L. limosa* and animal cells by *Arcobacter* might be conserved at the molecular level. This adds to recent evidence that 'virulence' genes, although initially described for pathogens, may in fact mediate both beneficial and pathogenic host-microbe interactions<sup>6</sup>.

To investigate whether *Arcobacter* might also engage beneficially with other Breviatea, we screened previously published<sup>1,4</sup> transcriptomic data of *P. bifurcata* and *Subulatomonas tetraspora* for genes affiliated with *Arcobacter*. In both cases, we identified so far unreported *Arcobacter* 16S rRNA gene sequences (Fig. 1b), as well as transcription of high-affinity, hydrogen-oxidizing Ni/Fe-hydrogenases (Extended Data Fig. 5). The phylogeny presented in Fig. 1b shows that the three Breviatea-associated *Arcobacter* do not form a monophyletic cluster, but are scattered phylogenetically among their animal-associated relatives. This could be interpreted in at least two ways. First, the symbiosis might have originated once in the ancestor of either phylum. Subsequently, *Arcobacter* co-diversified together with its hosts, and radiated to the other phylum multiple times, potentially via intermediate, free-living forms. Alternatively, symbiotic *Arcobacter* might have evolved from free-living forms multiple times. In either case, this resulted in the use of a similar molecular mechanism for colonization (see above). To assess the co-occurrence of Breviatea and *Arcobacter* in present-day marine sediments, currently available shotgun metagenomes were screened for the presence or absence of *L. limosa* and other Breviatea, as well as for Breviatea-associated *Arcobacter*. Thirteen out of 25 samples obtained from sediment horizons favourable for growth of Breviatea/*Arcobacter* consortia potentially contained Breviatea ( $P < 0.01$ ), and 17 potentially contained Breviatea-associated *Arcobacter*, including all samples positive for Breviatea (Extended Data Table 1). Although limited metagenomic data are available at present, they are consistent with the ecological interaction of Breviatea and *Arcobacter* observed for the enriched consortia. The evolutionary roots of this interaction can only be resolved by future investigation of more examples of Breviatea/*Arcobacter* consortia.

In conclusion, we have shown that *L. limosa* is a newly identified anaerobic protist, colonized by hydrogen-oxidizing *Arcobacter*. This colonization provides benefits to both partners via interspecies



hydrogen transfer, which enables the activity of a newly identified NADH-dependent fusion hydrogenase in *L. limosa*, leading to increased ATP yield. The molecular mechanism of colonization may involve specific interactions between *L. limosa* fibronectin type III domain-containing proteins and *Arcobacter* proteins similar to 'virulence' factors previously shown to mediate infection of animal tissue by related bacteria. The detection of *Arcobacter* genes in transcriptomes of other Breviatea shows that these protists probably engage in similar symbioses.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 3 August 2015; accepted 5 May 2016.**

**Published online 1 June 2016.**

1. Brown, M. W. *et al.* Phylogenomics demonstrates that breviate flagellates are related to opisthokonts and apusomonads. *Proc. R. Soc. B* **280**, 20131755 (2013).
2. Parfrey, L. W., Lahr, D. J. G., Knoll, A. H. & Katz, L. A. Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proc. Natl Acad. Sci. USA* **108**, 13624–13629 (2011).
3. Planavsky, N. J. *et al.* Low mid-Proterozoic atmospheric oxygen levels and the delayed rise of animals. *Science* **346**, 635–638 (2014).
4. Stairs, C. W. *et al.* A SUF Fe-S cluster biogenesis system in the mitochondrion-related organelles of the anaerobic protist *Pygmaea*. *Curr. Biol.* **24**, 1176–1186 (2014).
5. Ferreira, S., Queiroz, J. A., Oleastro, M. & Domingues, F. C. Insights in the pathogenesis and resistance of *Arcobacter*: a review. *Crit. Rev. Microbiol.* **42**, 364–383 (2016).
6. Sayavedra, L. *et al.* Abundant toxin-related genes in the genomes of beneficial symbionts from deep-sea hydrothermal vent mussels. *eLife* **4**, e07966 (2015).
7. McFall-Ngai, M. *et al.* Animals in a bacterial world, a new imperative for the life sciences. *Proc. Natl Acad. Sci. USA* **110**, 3229–3236 (2013).
8. Suga, H. *et al.* The *Capsaspora* genome reveals a complex unicellular prehistory of animals. *Nature Commun.* **4**, 2325 (2013).
9. Merchant, S. S. *et al.* The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* **318**, 245–250 (2007).
10. Stairs, C. W., Leger, M. M. & Roger, A. J. Diversity and origins of anaerobic metabolism in mitochondria and related organelles. *Phil. Trans. R. Soc. B* **370**, 20140326 (2015).
11. Schut, G. J. & Adams, M. W. W. The iron-hydrogenase of *Thermotoga maritima* utilizes ferredoxin and NADH synergistically: a new perspective on anaerobic hydrogen production. *J. Bacteriol.* **191**, 4451–4457 (2009).
12. Hrdy, I. *et al.* *Trichomonas* hydrogenosomes contain the NADH dehydrogenase module of mitochondrial complex I. *Nature* **432**, 618–622 (2004).
13. Stams, A. J. & Plugge, C. M. Electron transfer in syntrophic communities of anaerobic bacteria and archaea. *Nature Rev. Microbiol.* **7**, 568–577 (2009).
14. Moser, I., Schroeder, W. & Salnikow, J. *Campylobacter jejuni* major outer membrane protein and a 59-kDa protein are involved in binding to fibronectin and INT 407 cell membranes. *FEMS Microbiol. Lett.* **157**, 233–238 (1997).
15. Monteville, M. R., Yoon, J. E. & Konkel, M. E. Maximal adherence and invasion of INT 407 cells by *Campylobacter jejuni* requires the CadF outer-membrane protein and microfilament reorganization. *Microbiology* **149**, 153–165 (2003).
16. van Alphen, L. B. *et al.* Active migration into the subcellular space precedes *Campylobacter jejuni* invasion of epithelial cells. *Cell. Microbiol.* **10**, 53–66 (2008).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank T. Hargesheimer, D. Liu, G. Klockgether, P. Hach, R. Appel and I. Kattelmann for technical assistance, A. Simpson, G. Strous and Fulvio Reggiori for comments on electron micrographs, and C. Hubert, E. Ruff, S. Ahmerkamp and N. Dubilier for discussions. This study was supported by European Research Council starting grant MASEM 242635 (M.S., E.H., J.C.), the Campus Alberta Innovation Chair Program (M.S., E.H., X.D.), the Canadian Foundation for Innovation (M.S.), the Alberta Small Equipment Grant Program (M.S.), the German Federal State Nordrhein-Westfalen (M.S.), the Max Planck Society, and the Natural Sciences and Engineering Research Council of Canada for a Banting fellowship to M.K. and a Discovery Grant to M.S.

**Author Contributions** E.H. and J.C. performed sampling, cultivation and physiological experiments. M.K. performed proteomics and data analysis. D.R. performed transmission electron microscopy. S.L. and E.H. performed scanning electron microscopy. E.H. performed CARD-FISH imaging. H.T. performed next-generation sequencing. H.G.-V. performed read processing, assembly and binning. E.H. performed *in silico* processing of next-generation sequencing data with assistance from H.G.-V., X.D. and M.S. M.W.B., C.W.S. and A.J.R. analysed sequences for *Arcobacter* associated with *S. tetraspora*. E.H., B.V. and K.B. performed chemical analysis with input from J.M., K.-U.H. and M.S. The experimental design was developed jointly by M.S., E.H., J.M., M.K. and K.-U.H. E.H. wrote the manuscript with input from all co-authors.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.S. ([mstros@ucalgary.ca](mailto:mstros@ucalgary.ca)) or E.H. ([emmohamann@gmail.com](mailto:emmohamann@gmail.com)).

## METHODS

**Cultivation of *L. limosa*.** Sediment samples for the initial enrichment of *L. limosa* were obtained from a tidal flat in the German Wadden Sea (53.73585° N, 7.69905° E). For the enrichments, 100 ml laboratory bottles (DURAN Glastechnik) were filled with 50 ml sediment from different depths (0.5–2 cm and 6–8 cm). The bottles were filled with cultivation medium and closed without creating a headspace. The cultivation medium was based on HEPES-buffered seawater (34 g l<sup>-1</sup>, 1 mM HEPES, pH 8) (Red Sea Deutschland) and contained prey bacteria for protists (final density approximately 10<sup>9</sup> cells per millilitre), as well as either nitrate, nitrite or nitrous oxide (0.2 mM final) as electron acceptor for bacterial denitrification. To allow for growth of micrometre-scaled protists, the small rod-shaped, strictly aerobic gammaproteobacterium *A. macleodii* (strain ATCC 27126, 0.6–0.8 µm × 1.4–2 µm in size) was selected as prey bacterium. Prey bacteria were grown on solid marine broth medium (ATCC medium 2216). The cells were washed three times by serial centrifugation in sterile seawater before they were added to the enrichment cultures. Enrichment cultures for protists were incubated in the dark at 22 °C and the nitrate as well as nitrite concentrations were regularly monitored. Depletion of electron acceptor was prevented by periodically adding small portions of anoxic concentrated stock solutions of nitrate, nitrite (from a 200 mM stock) or pure N<sub>2</sub>O to the cultures. Depletion of food was prevented by adding portions of anoxic suspensions of prey bacteria to the cultures. Growth of protists was regularly inspected by light microscopy. When enrichment was observed, a subsample of culture was transferred to sediment free anoxic medium for further investigations. To best preserve the microbial community naturally associated with *L. limosa*, we did not isolate single protists. Instead we performed serial transfers into fresh medium until the presence of only one species could be confirmed via Sanger sequencing of 18S rDNA genes and light microscopy. For all results presented in this study we only used one strain (strain LL-12) obtained from an enrichment culture with nitrous oxide. This strain was maintained in medium containing 2 mM nitrous oxide and transferred weekly. For the cultivation of *L. limosa* in the absence of *Arcobacter*, we grew a subculture in medium containing no nitrous oxide. This culture was transferred once before proteomic analysis (see below).

**Cultivation of *Arcobacter*.** For the cultivation of *Arcobacter* in absence of *L. limosa*, we first incubated *L. limosa*/*Arcobacter* consortia in anoxic medium containing dissolved nutrients but no prey bacteria. The cultivation medium was based on HEPES-buffered seawater (34 g l<sup>-1</sup>, 1 mM HEPES, pH 8) and supplemented with the following compounds: sodium acetate (10 mM), yeast extract (0.1 g l<sup>-1</sup>), sodium phosphate (2 mM) and ammonium chloride (2 mM). The medium was filled in air-tight cultivation bottles and made anoxic by flushing it with a mixture of 5% H<sub>2</sub>, 10% CO<sub>2</sub> and 85% N<sub>2</sub>. Afterwards, part of the culture headspace was replaced with nitrous oxide to achieve a nitrous oxide concentration of 5 mM in the liquid. Under the given conditions no growth of *L. limosa* could be observed. The enrichment cultures were transferred twice before proteomic analysis (see below).

**Physiological experiments.** Several growth experiments were performed to evaluate the importance of bacterial hydrogen oxidation for the fitness of *L. limosa* (Fig. 3 and Extended Data Fig. 6). As electron acceptor for hydrogen oxidation we used nitrous oxide, which is the last intermediate in the serial denitrification pathway. Nitrous oxide is non-toxic and its reduction does not produce any intermediates. Thus, it allows for an easy and unambiguous determination of denitrification rates by measuring the consumption of nitrous oxide in the medium. To compare the growth efficiency of *L. limosa* under different conditions, we determined cell numbers, growth yields as well as respiration rates. To determine cell numbers of *L. limosa*, 50 µl subsamples of liquid culture were mixed with formaldehyde solution (0.02% final), which immobilized swimming cells. The cell abundance was determined by manually counting with an improved Neubauer counting chamber (BRAND counting chamber, Neubauer improved, 0.1 mm depth). For bacterial cell counts, cells were fixed in 1.8% formaldehyde solution (prepared in 34 g l<sup>-1</sup> sterile seawater, 1 mM HEPES, pH 8) for 2 h at room temperature (~20 °C), filtered onto 0.22 µm white polycarbonate filters (supported by 0.45 µm cellulose nitrate filter, Whatman) and stained with the DNA-specific stain DAPI (1 µg ml<sup>-1</sup>, 3 min, at 37 °C). The filters were evaluated by epifluorescence microscopy (Leila DM: Oram centra mercury-vapour lamp) and the cells were manually counted at ×1,000 magnification.

Growth yields were calculated by measuring the bacterial cell numbers as well as the eukaryotic cell numbers between two different time points during exponential growth. The differences in bacterial abundance divided by the difference in eukaryotic cell numbers provided a proxy for the gross growth efficiency.

Respiration rates were determined by adding <sup>13</sup>C-enriched *A. macleodii* bacteria to an exponentially growing culture of *L. limosa*. Digestion of such labelled cells leads to the production of <sup>13</sup>C bicarbonate, which was measured after conversion to <sup>13</sup>CO<sub>2</sub> (see Chemical analysis). To investigate the growth of *L. limosa* in the presence of antibacterial antibiotics, ampicillin and streptomycin (0.1 mg ml<sup>-1</sup> each) were added to the medium. Since these antibiotics act on growing cells, they

did not affect the availability of food bacteria, which were provided in excess at the start of the incubation.

To inhibit denitrifying activity, we added 5% acetylene final to the headspace of the cultures and dissolved it by gently shaking the cultures. Acetylene inhibits nitrous oxide reductase, the enzyme complex that mediates nitrous oxide respiration.

To test the effect of nitrate on the growth of *L. limosa*, we added 2 mM sodium nitrate to the cultivation medium. In addition, we incubated a culture in medium that contained 0.2 mM dissolved oxygen. Cell numbers of *L. limosa* in such treated cultures were compared with a culture provided with nitrous oxide (2.2 mM) and with a control culture that did not contain an electron acceptor for hydrogen oxidation.

**Chemical analysis.** Nitrous oxide and hydrogen was measured from the gas headspace of the cultures using a GAM 400 mass spectrometer (In Process Instruments). Volatile fatty acids were measured with a Syham HPLC system (Fürstendfeldbruck) equipped with an Aminex HPX-87 H HPLC column (300 × 7.8 mm) and 5 mM H<sub>2</sub>SO<sub>4</sub> as eluent. Separation was performed in isothermal mode at 40 °C and the eluted compounds were simultaneously detected with an ultraviolet and a refractive index detector at a detection limit of 0.1 mM. As calibration standard, a mixture of the fatty acids succinate, lactate, formate, acetate, propionate and butyrate was measured at different concentrations.

Respiration rates were measured as dissolved inorganic <sup>13</sup>C released by the digestion of <sup>13</sup>C-enriched prey bacteria. Subsamples (1 ml) were poisoned with 0.2% zinc acetate solution at different time-points during the experiments. The isotopic component of dissolved inorganic <sup>13</sup>C was determined after acidifying with hypo-phosphoric acid (1% final) and analysed on a gas chromatography–isotope ratio monitoring mass spectrometer (Optima Micromass). To convert isotopic compositions to concentrations (moles per litre), standard solutions of NaHCO<sub>3</sub> with known dissolved inorganic <sup>13</sup>C concentrations were measured.

**CARD-FISH.** CARD-FISH was performed on polycarbonate filters as described elsewhere<sup>17</sup>. For the hybridization of ribosomal RNA we used the following probes. For eukaryotes, Euk516, ACCAGACTTGCCCTCC (5′–3′, 0% formamide); for Epsilonproteobacteria, Epsy914, GGTCCTCGTCTATTCTT (5′–3′, 35% formamide); and for Alteromonas, Alt184, CCCGTTTGGTCCGAAGAC (5′–3′, 25% formamide). Before microscopic evaluation, all samples were counter-stained with DAPI and embedded in a 3:1 mixture of Citifluor-Vectashield (Citifluor/Vector Labs). Cells were imaged at ×1,000 magnification with an epifluorescence microscope (AxioSkop 2 MOT Plus, Carl Zeiss) connected to an AxioCam MRm camera (Carl Zeiss).

**Electron microscopy.** For transmission electron microscopy, cells were harvested at 2,000 r.p.m. (260g) using a Stat Spin Microprep 2 table-top centrifuge. After centrifugation, the pellet was vitrified in a BAL-TEC HPM-010 high-pressure freezer. The samples were substituted at –90 °C in a solution containing 0.1% tannic acid and 0.5% glutaraldehyde in anhydrous acetone for 72 h, and for additional 8 h in 2% OsO<sub>4</sub> in anhydrous acetone. After a further incubation over 20 h at –20 °C, samples were warmed up to +4 °C and washed with anhydrous acetone. The samples were embedded at room temperature in Agar 100 (Epon 812 equivalent) at 60 °C over 24 h. After ultrathin sectioning (60 nm), sections were counter-stained with lead citrate. Samples were analysed with a Philips CM 120 transmission electron microscope (Philips) and images were taken with a TemCam F416 CMOS camera (TVIPS).

For scanning electron microscopy, cells were harvested at 400 r.p.m. for 6 min using a table-top centrifuge. After centrifugation, the pellet was placed on Teflon slides. We then waited 5 min to allow attachment of cells to the slide and fixed them with 2% glutaraldehyde solution (in 34 g l<sup>-1</sup> sterile seawater, 1 mM HEPES, pH 8) for 60 min at room temperature. Fixation was followed by a washing step in MilliQ-water and an ethanol dehydration series in 30%, 50%, 70%, 90% and 100% ethanol (20 min each). Finally, the specimens were subjected to critical-point drying with CO<sub>2</sub> to remove any volatile solvents. The objects were stored in a silica-filled desiccator until microscopic evaluation. Imaging was performed with a Nova NanoLab 600 scanning electron microscope (FEI). For better identification of bacteria (in Fig. 1c), colours were manually added and the background was removed using the imaging processing software GNU Image Manipulation Program (the GIMP team, version 2.8.14).

**Percoll density-gradient centrifugation.** Enrichment of *L. limosa*/*Arcobacter* consortia from suspended bacterial cells was done by Percoll (Sigma-Aldrich) -based density-gradient centrifugation. To form a density gradient, 4.5 ml Percoll was mixed with 3.5 ml buffered sterile seawater (1 mM HEPES), filled in 10 ml centrifugation tubes and centrifuged for 30 min at 10,000g and 15 °C. Afterwards, 1 ml of culture was carefully loaded on top of the gradient and immediately centrifuged for another 10 min at 8,000g at 15 °C. Cells of *L. limosa* were enriched in a dim white band in the upper third of the gradient, approximately 0.5 cm above the bacteria. The *L. limosa* fraction was diluted in 50 ml sterile seawater and washed

twice by serial centrifugation at 800g for 5 min (15 °C) to remove Percoll and suspended bacteria. The enriched *L. limosa* fraction was pelleted, frozen and stored at −80 °C until DNA extraction.

**Next-generation sequencing and *in silico* procedures.** Shotgun metagenomic Illumina libraries were constructed from genomic DNA of enriched *L. limosa*/*Arcobacter* consortia (see Percoll density-gradient centrifugation). Sequencing yielded ~7.5 million raw 2 × 250 bp paired-end MiSeq reads. Reads were quality trimmed to Q10 and filtered to a minimum length of 99 nucleotides using Neson (https://github.com/Victorian-Bioinformatics-Consortium/nesoni). Initial assembly of the quality-filtered reads was performed with Velvet<sup>18</sup>. Bacterial genomes were then binned on the basis of percentage GC content, tetranucleotide composition and sequence coverage using Metawatt<sup>19</sup>. The binning yielded provisional whole-genome sequences for *L. limosa*, *Arcobacter* as well as for co-enriched *Lacinutrix* and *Colwellia*. For a refined assembly of the *L. limosa* genome, we filtered out all reads that mapped to the bacterial bins using Bowtie2 (ref. 20). The remaining reads were then used to generate a final assembly for *L. limosa*, using Velvet and Gapfiller<sup>21</sup>. This assembly was 48.1 Mb in size and had an N50 value of 9.5 kb. The length of the longest scaffold was 135.5 kb. Structural annotations and gene predictions for *L. limosa* were performed with the MAKER pipeline<sup>22</sup>. After an initial *ab initio* gene prediction with GeneMark-ES<sup>23</sup>, we refined the obtained gene models with evidence-driven gene predictions using Snap<sup>24</sup> and 48,530 assembled transcripts of a polyA-tail enriched transcriptome (see below). Repeat identification, annotation and masking were implemented by RepeatMasker<sup>25</sup>. Genome completeness was estimated by representation of 159 universal eukaryotic genes previously identified in *P. bifurca*<sup>1</sup> using hidden Markov model-based searches with hmmer and blast homology searches. Functional annotation was performed with the KEGG automatic annotation server<sup>26</sup>. For the core metabolism of *L. limosa*, we manually validated the correct prediction and annotation of each gene model by blastx and blastp homology searches against the UniProtKB/Swiss-Prot protein database. Translocation of gene products to mitochondria-related organelles was predicted on the basis of the presence of N-terminal target peptides using TargetP<sup>27</sup> and MitoPROT<sup>28</sup>. Transcriptional activity of genes was evaluated by mapping the reads of a polyA+ tail-enriched transcriptome to the predicted nucleotide gene models. Protein architectures and conserved protein domains were identified using HMMER<sup>29</sup>, the Pfam<sup>30</sup> database and the SMART<sup>31</sup> protein domain detection tools.

For the *Arcobacter* genome, gene predictions and functional annotations were performed using the online annotation-pipeline RAST<sup>32</sup>. Completeness of the *Arcobacter* genome was evaluated with CheckM<sup>33</sup> using the implemented set of conserved Epsilonproteobacteria reference genes. For the core metabolism and virulence related genes, we manually validated the correct prediction and annotation of each gene model with blastx and blastp homology searches against the UniProtKB/Swiss-Prot protein database.

For mRNA sequencing, RNA was preserved in RNAlater and was extracted as previously described<sup>34</sup>. Extracted RNA was treated with RQ1 DNase (Promega), purified with RNeasy MinElute columns (Qiagen) and stored in TE buffer at −80 °C. Approximately 1 µg of total RNA was used for preparation of an mRNA sequencing library following the Illumina TruSeq RNA Sample Preparation version 2 guide, using poly-T oligo-attached magnetic beads to enrich eukaryotic mRNA. The RNA library was sequenced on a MiSeq instrument in a 2 × 250 bp paired-end run.

**Phylogenetic analysis.** Eukaryotic phylogenies were determined with an alignment consisting of 16 universal eukaryotic genes from 88 different taxa covering all major eukaryotic groups as previously published<sup>35</sup>. The alignment was complemented with genes from *L. limosa*, *Pygusia bifurca*, *Breviata anathema* as well as *S. tetraspora*. Alignments were calculated with MAFFT<sup>36</sup> and phylogenies were constructed with RaxML<sup>37</sup> using the GTR+GAMMA model for the SSU-rDNA partition and the WAG replacement matrix with maximum likelihood estimated base frequencies for the amino-acid partitions. We performed 400 rapid bootstrap iterations followed by a search for the best-scoring maximum likelihood tree.

*Arcobacter* phylogenies were calculated using 16S rRNA sequences as phylogenetic marker genes. Alignments were constructed with MAFFT. Phylogenetic tree calculation was performed by bayesian inference using the software MrBayes<sup>38</sup> with the GTR substitution model and gamma rate variation.

Phylogenetic analysis for the putative NAD(P)H-dependent Fe-hydrogenase was done individually for the NAD(P)H-accepting domain and the Fe-hydrogenase domain. For this, closely related sequences were obtained from the NCBI non-redundant protein database and used to construct a sequence alignment with MAFFT. Phylogenies were inferred using RaxML and the WAG replacement matrix with fixed base frequencies. We ran 600 rapid bootstrap iterations followed by a search for the best scoring maximum likelihood tree.

**Protein extraction and peptide preparation.** For proteomics, three parallel cultures of *L. limosa* were grown with nitrous oxide in the presence of *Arcobacter* and

three cultures were grown without nitrous oxide and in the absence of *Arcobacter*. In addition, three cultures of *Arcobacter* were grown with nitrous oxide and in the absence of *L. limosa* (see Cultivation for details on medium composition). The cultures were harvested during the early exponential growth phase and immediately frozen at −80 °C.

For each of the three treatments, we prepared tryptic digests from three biological replicates following the filter-aided sample preparation protocol described in ref. 39. In brief, SDT-lysis buffer (4% (w/v) SDS, 100 mM Tris-HCl pH 7.6, 0.1 M DTT) was added in a 1:10 sample:buffer ratio to the sample pellets. Samples were heated to 95 °C for 10 min followed by pelleting of debris for 5 min at 21,000g. Thirty microlitres of the cleared lysate were mixed with 200 µl of UA solution (8 M urea in 0.1 M Tris/HCl pH 8.5) in a 10 kDa MWCO 500 µl centrifugal filter unit (VWR International) and centrifuged at 14,000g for 40 min. Two hundred microlitres of UA solution were added again and centrifugal filter spun at 14,000g for 40 min. One hundred microlitres of IAA solution (0.05 M iodoacetamide in UA solution) were added to the filter and incubated at 22 °C for 20 min. The IAA solution was removed by centrifugation and the filter was washed three times by adding 100 µl of UA solution and then centrifuging. The buffer on the filter was then changed to ABC (50 mM ammonium bicarbonate), by washing the filter three times with 100 µl of ABC. Two micrograms of MS grade trypsin (Thermo Scientific Pierce) in 40 µl of ABC were added to the filter, and filters were incubated overnight in a wet chamber at 37 °C. The next day, peptides were eluted by centrifugation at 14,000g for 20 min, followed by addition of 50 µl of 0.5 M NaCl and again centrifugation. Peptides were de-salted using C18 spin columns (Thermo Scientific Pierce) according to the manufacturer's instructions. Approximate peptide concentrations were determined using a Pierce Micro BCA assay (Thermo Scientific Pierce).

**One-dimensional liquid chromatography–tandem mass spectrometry.** Samples were analysed by one-dimensional liquid chromatography–tandem mass spectrometry (LC–MS/MS) using a block-randomized design as previously described<sup>40</sup>. Two blank runs were done between samples to reduce carry over. For each sample, a technical replicate was run. For each run, 800 ng of peptide were loaded onto a 2 cm, 75 µm ID C18 Acclaim PepMap 100 pre-column (Thermo Fisher Scientific) using an EASY-nLC 1000 Liquid Chromatograph (Thermo Fisher Scientific) set up in two-column mode. The pre-column was connected to a 50 cm × 75 µm analytical EASY-Spray column packed with PepMap RSLC C18, 2 µm material (Thermo Fisher Scientific), which was heated to 35 °C via the integrated heating module. The analytical column was connected via an Easy-Spray source to a Q Exactive Plus Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific). Peptides were separated on the analytical column at a flow rate of 225 nl min<sup>−1</sup> using a 260 min gradient going from buffer A (0.2% formic acid, 5% acetonitrile) to 20% buffer B (0.2% formic acid in acetonitrile) in 200 min, then from 20% to 35% buffer B in 40 min and ending with 20 min at 100% buffer B. Eluting peptides were ionized with electrospray ionization and analysed in Q Exactive Plus. Full scans were acquired in the Orbitrap mass spectrometer at 70,000 resolution. MS/MS scans of the 15 most abundant precursor ions were acquired in the Orbitrap mass spectrometer at 17,500 resolution. The mass (*m/z*) 445.12003 was used as lock mass as described in ref. 41 with the modification that lock mass was detected in the full scan rather than by separate selected ion monitoring scan injection. Lock mass use was set to 'best'. Ions with charge state +1 were excluded from MS/MS analysis. Dynamic exclusion was set to 30 s. Roughly 160,000 MS/MS spectra were acquired per sample run.

**Protein identification, quantification and statistics.** For protein identification, a custom protein sequence database containing 28,966 proteins, predicted from the *L. limosa*, *Arcobacter* sp., *Alteromonas*, *Colwellia* and *Lacinutrix* provisional whole-genome sequences, was used. The database was submitted to the PRIDE repository (see below). For protein identification, technical replicates were combined and MS/MS spectra were searched against the database using the Sequest HT node in Proteome Discoverer version 2.0.0.802 (Thermo Fisher Scientific) with the following parameters: trypsin (full), maximum two missed cleavages, 10 p.p.m. precursor mass tolerance, 0.6 Da fragment mass tolerance and maximum three equal dynamic modifications per peptide. The following three dynamic modifications were considered: oxidation on M (+15.995 Da), carbamidomethyl on C (+57.021 Da) and acetyl on protein N terminus (+42.011 Da). False discovery rates (FDRs) for peptide spectral matches (PSMs) were calculated and filtered using the Percolator Node in Proteome Discoverer. The Percolator algorithm<sup>42</sup> 'uses semi-supervised learning and a decoy database search strategy to learn to distinguish between correct and incorrect PSMs'. Percolator was run with the following settings: maximum delta Cn 0.05, a strict target FDR of 0.01, a relaxed target FDR of 0.05 and validation based on *q* value.

Search results for all samples were combined into a multiconsensus report with Proteome Discoverer and additional filtering criteria applied on the protein level, which were at least six PSMs per protein, at least one unique peptide and only PSMs with a concatenated rank of 1. This resulted in the following overall FDRs:



0.6% for PSMs, 2.6% for peptides and 5% for proteins. The multiconsensus report was then exported as a tab-delimited file for further processing.

For protein quantification, normalized spectral abundance factors (NSAFs) were calculated on the basis of PSMs using the method described previously<sup>43</sup> and multiplied by 10,000. The NSAF  $\times$  10,000 value gives the relative abundance of a protein in a sample as a fraction of 10,000. For statistical analyses of differences between treatments, *L. limosa* and *Arcobacter* proteins were analysed separately: that is, separate tables were generated. For both organisms the set of proteins was reduced by only including proteins that had at least three NSAF values greater than 0 across all nine samples (three treatments  $\times$  three biological replicates) and at least one NSAF value greater than 0 in each of the treatments that were compared in the respective statistical test. The NSAF  $\times$  10,000 values in these reduced tables were then again normalized to 10,000 for each sample to generate organism-specific NSAF values (orgNSAFs) as described in ref. 44. This normalization procedure ensures that differences in organism abundance across samples do not lead to the false detection of differentially expressed proteins. The tables with orgNSAFs were loaded into Perseus software (version 1.5.1.6, <http://www.perseus-framework.org/doku.php>) and  $\log_2$  transformed. Missing values produced by  $\log_2(0)$  were replaced by sampling from a normal distribution assuming that the missing values were on the lower end of abundance (normal distribution parameters in Perseus: width 0.3, down shift 1.8, do separately for each column). A *t*-test with permutation-based FDR calculation was used to detect proteins that differed significantly in their expression level between two treatments. The statistical method implemented in Perseus that we used was based on the 'significance analysis of microarrays' described in ref. 45, which by using a permutation-based FDR, accounted for the multiple-testing problem inherent in testing for significant expression differences for a large number of genes. The following parameters were used for the test: groupings were not preserved for randomizations, both sides, 250 randomizations, FDR of 1% and  $s_0$  of 0.1.

No statistical methods were used to predetermine sample size. The investigators were not blinded to allocation during experiments and outcome assessment. The experiments were not randomized except for the proteomics approach (see section on one-dimensional liquid chromatography–tandem mass spectrometry).

**Screening of sediment metagenomes.** A SRAdB R/Bioconductor<sup>46</sup> query of the NCBI Sequence Read Archive (SRA) with search term 'marine sediment' yielded 130 shotgun metagenomes. After downloading and converting the SRA files to FASTQ format, the FASTQ short, <300 bp reads were mapped onto the assembled contigs (only those >5 kbp) of *L. limosa* and *Arcobacter*, using BMap with parameters 'ambiguous=random qtrim=lr trimq=10 minid=0.90'. The remaining long, >300 bp reads were quality trimmed using BMap with parameters 'qtrim=lr trimq=10' before mapping onto the same contigs using BWASW<sup>47</sup> with default settings. For robustness of phylogenetic classification, the low-complexity reads identified using sga<sup>48</sup> with parameters 'preprocess --dust --dust-threshold=4' were filtered out from the mapped reads; with diamond blastx<sup>49</sup> the remaining mapped reads were searched against a database containing amino-acid sequences of genus-level (Bacteria) or family-level (Eukarya) representatives of all forms of life, including *L. limosa* and other Breviatea, as well as *L. limosa*-affiliated *Arcobacter*, *A. nitrofigilis* and seven other *Arcobacter*. Only reads with hits to Breviatea, *L. limosa*-affiliated *Arcobacter* and *A. nitrofigilis*, and only hits within 10% of the best bit-score with *e*-value lower than 0.01 and min bit-score at least 50, were considered. Distribution of positive hits over multiple loci of *L. limosa* and *Arcobacter* sp. genome sequences was verified manually. Of the 130 available metagenomes, 15 had to be rejected because they were not obtained from marine sediments, and 65 were rejected because they contained too few reads to enable detection of *L. limosa*. Of the remaining 50 metagenomes, 25 were obtained from the top of the sediment, a potentially favourable habitat for *L. limosa*. The remaining 25 were obtained from deeper sediment horizons with probably unfavourable conditions. For the latter, we calculated an average false positive rate for Breviatea detection of 0.004 reads per 1 million reads. This false positive rate is probably overestimated, because we cannot completely exclude that the metagenomes from unfavourable conditions did not contain Breviatea or were contaminated during sampling. The false positive rate was used to calculate, for each sample, using a binomial distribution, the probability that the actual number of reads assigned to Breviatea was coincidental and not related to the potential presence of Breviatea. We did not apply the same statistical procedure to the read maps for *Arcobacter*, because classifying samples as unfavourable for *Arcobacter* was not possible since *Arcobacter* can occupy many different ecological niches. Extended Data Table 1 presents the results and SRA accession numbers for the analysed metagenomes.

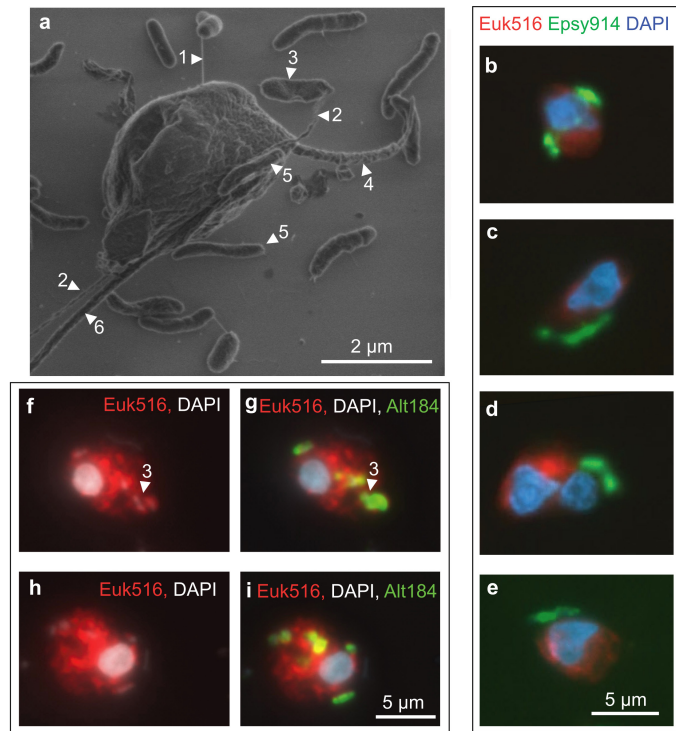
**Bioenergetics and hydrogen fluxes.** Per-cell hydrogen production rates were calculated from nitrous oxide consumption rates and cell numbers presented in Fig. 3, assuming a stoichiometry of 1:3 mol H<sub>2</sub>:mol N<sub>2</sub>O (Fig. 2). The hydrogen concentration sustained in the mitochondria-related organelles of *L. limosa* by a single *Arcobacter* epibiont was calculated to be 10  $\mu$ M with the equation

$F \text{ (mol H}_2\text{ s}^{-1}) = D/d \times C \times A$ , with *F*, the hydrogen flux ( $\sim 9 \times 10^{-18}$  mol s<sup>-1</sup>); *D*, the diffusivity of H<sub>2</sub> in water ( $\sim 4.6 \times 10^{-9}$  m<sup>2</sup> s<sup>-1</sup>); *d*, the distance between the H<sub>2</sub> source and the H<sub>2</sub> sink ( $2.1 \times 10^{-6}$  m); *C*, the H<sub>2</sub> concentration in the mitochondria-related organelles (mM or mol m<sup>-3</sup>); and *A*, the area of diffusion ( $0.4 \times 10^{-12}$  m<sup>2</sup>). The maximum hydrogen concentration enabling the reaction  $\text{NADH} + \text{H}^+ = \text{NAD}^+ + \text{H}_2$  ( $\Delta G = +18$  kJ mol<sup>-1</sup> (ref. 13)) was calculated to be 5.4  $\mu$ M (with Henry's law constant for H<sub>2</sub> =  $6.33 \times 10^9$  Pa and an NADH/NAD<sup>+</sup> ratio of 10 (ref. 50)). The calculations are presented in Supplementary Table 2.

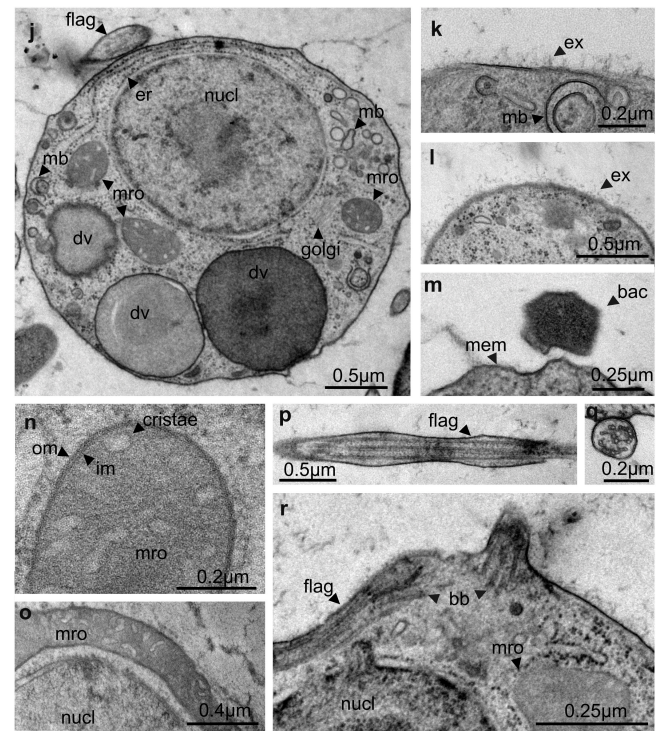
**Reference material and data availability.** The culture was deposited in the cryo-preserved state in liquid nitrogen at the American Type Culture Collection and has been accessioned as *Lenisia limosa* (strain LL-12) AcqID-00721, and will be publicly available as soon as possible. In the meantime, we will provide *L. limosa* from our laboratory stocks upon request. Sequence data are available for download from the NCBI SRA database and the Whole Genome Shotgun database and are grouped as NCBI BioProject PRJNA277740. The SSU rRNA gene for *L. limosa* is available in GenBank under accession number KT023596. The mass spectrometry proteomics data and the protein sequence database have been deposited in the ProteomeXchange Consortium via the PRIDE partner repository with the data set identifier PXD003275.

- Pernthaler, A., Pernthaler, J. & Amann, R. Fluorescence in situ hybridization and catalyzed reporter deposition for the identification of marine bacteria. *Appl. Environ. Microbiol.* **68**, 3094–3101 (2002).
- Zerbino, D. R. & Birney, E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* **18**, 821–829 (2008).
- Strous, M., Kraft, B., Bisdorf, R. & Tegetmeyer, H. E. The binning of metagenomic contigs for microbial physiology of mixed cultures. *Front. Microbiol.* **3**, 410 (2012).
- Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357–359 (2012).
- Nadalín, F., Vezzi, F. & Policriti, A. GapFiller: a de novo assembly approach to fill the gap within paired reads. *BMC Bioinformatics* **13** (Suppl. 14), S8 (2012).
- Cantarel, B. L. et al. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* **18**, 188–196 (2008).
- Ter-Hovhannissyan, V., Lomsadze, A., Chernoff, Y. O. & Borodovsky, M. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Res.* **18**, 1979–1990 (2008).
- Korf, I. Gene finding in novel genomes. *BMC Bioinformatics* **5**, 59 (2004).
- Chen, N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinformatics* **25**, 4.10.1–4.10.14 (2004).
- Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. & Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* **40**, D109–D114 (2012).
- Emanuelsson, O., Nielsen, H., Brunak, S. & von Heijne, G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* **300**, 1005–1016 (2000).
- Claros, M. G. MitoProt, a Macintosh application for studying mitochondrial proteins. *Comput. Appl. Biosci.* **11**, 441–447 (1995).
- Finn, R. D., Clements, J. & Eddy, S. R. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* **39**, W29–W37 (2011).
- Finn, R. D. et al. The Pfam protein families database. *Nucleic Acids Res.* **38**, D211–D222 (2010).
- Letunic, I., Doerks, T. & Bork, P. SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res.* **43**, D257–D260 (2015).
- Glass, E. M., Wilkening, J., Wilke, A. & Antonopoulos, D. & Meyer, F. Using the metagenomics RAST server (MG-RAST) for analyzing shotgun metagenomes. *Cold Spring Harb. Protoc.* **5**, <http://dx.doi.org/10.1101/pdb.prot5368> (2010).
- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *PeerJ* **3**, 1–4 (2014).
- Smith, C. J., Nedwell, D. B., Dong, L. F. & Osborn, A. M. Diversity and abundance of nitrate reductase genes (*narG* and *napA*), nitrite reductase genes (*nirS* and *nrfA*), and their transcripts in estuarine sediments. *Appl. Environ. Microbiol.* **73**, 3612–3622 (2007).
- Parfrey, L. W. et al. Broadly sampled multigenic analyses yield a well-resolved eukaryotic tree of life. *Syst. Biol.* **59**, 518–533 (2010).
- Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
- Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
- Ronquist, F. et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).
- Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nature Methods* **6**, 359–362 (2009).
- Oberg, A. L. & Vitek, O. Statistical design of quantitative mass spectrometry-based proteomic experiments. *J. Proteome Res.* **8**, 2144–2156 (2009).
- Olsen, J. V. et al. Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell. Proteomics* **4**, 2010–2021 (2005).

42. Spivak, M., Weston, J., Bottou, L., Käll, L. & Noble, W. S. Improvements to the percolator algorithm for peptide identification from shotgun proteomics data sets. *J. Proteome Res.* **8**, 3737–3745 (2009).
43. Florens, L. *et al.* Analyzing chromatin remodeling complexes using shotgun proteomics and normalized spectral abundance factors. *Methods* **40**, 303–311 (2006).
44. Mueller, R. S. *et al.* Ecological distribution and population physiology defined by proteomics in a natural microbial community. *Mol. Syst. Biol.* **6**, 374 (2010).
45. Tusher, V. G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl Acad. Sci. USA* **98**, 5116–5121 (2001).
46. Zhu, Y., Stephens, R. M., Meltzer, P. S. & Davis, S. R. SRAdb: query and use public next-generation sequencing data from within R. *BMC Bioinformatics* **14**, 19 (2013).
47. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
48. Simpson, J. T. & Durbin, R. Efficient de novo assembly of large genomes using compressed data structures. *Genome Res.* **22**, 549–556 (2012).
49. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nature Methods* **12**, 59–60 (2015).
50. Hung, Y. P., Albeck, J. G., Tantama, M. & Yellen, G. Imaging cytosolic NADH-NAD<sup>+</sup> redox state with a genetically encoded fluorescent biosensor. *Cell Metab.* **14**, 545–554 (2011).

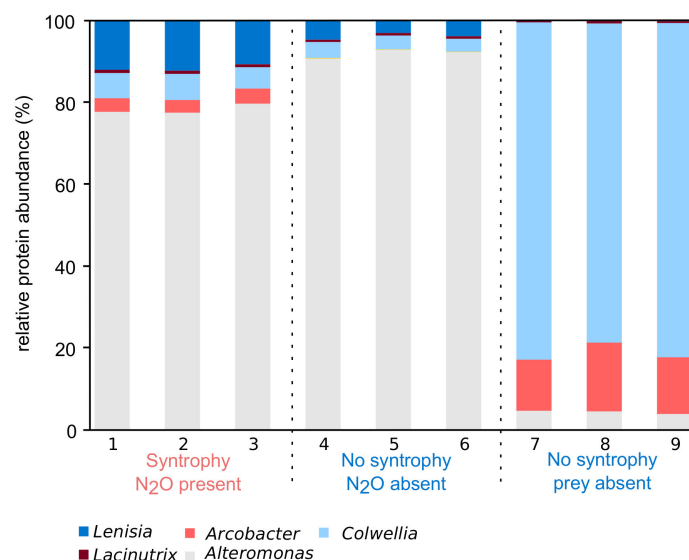


**Extended Data Figure 1 | Micrographs for *L. limosa* and epibiotic *Arcobacter*.** **a**, Scanning electron micrograph showing *L. limosa* and associated bacteria. Pilus (1) connecting *Arcobacter* (5) with *L. limosa*. Pseudopodial extensions (2) are used for the acquisition of prey bacteria (3) (*Alteromonas*). Short anterior flagellum (4). Long posterior flagellum (6). **b–e**, CARD-FISH labelling with probes targeting the SSU rRNA of *L. limosa* (Euk516 in red) and *Arcobacter* (Epsy914 in green). The scale bar applies to all figures. **f–i**, CARD-FISH labelling of *L. limosa* with probes



targeting the SSU rRNA of *A. macleodii* (Alt184 in green). The scale bar applies to all figures. **j–r**, Transmission electron micrographs showing different structural features of *L. limosa*. Mitochondria-related organelle (mro), nucleus (nucl), digestive vacuoles (dv), double basal body (bb), endoplasmic reticulum (er), inner (im) and outer membrane (om), tubular cristae (cristae), extracellular matrix (ex), bacterium (bac), membrane (mem), flagellum (flag), multivesicular body (mb). For **a–i**, each specimen shown represents at least ten specimens for which images were recorded.



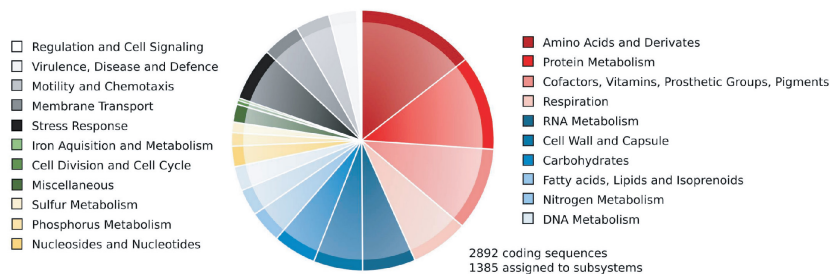


**Extended Data Figure 2 | Relative abundance of *L. limosa* and co-enriched bacteria under different growth conditions.** The abundance of *L. limosa* and its associated microbiota was determined at three different conditions (treatments) with three independent experiments per treatment: 1–3, presence of nitrous oxide and prey bacteria;

4–6, absence of nitrous oxide and presence of prey bacteria; 7–9, presence of nitrous oxide, dissolved organic nutrients and hydrogen and absence of prey bacteria. Relative abundances were determined via proteomics and estimated on the basis of the total normalized spectrum count per population.

**The 3.0 Mb genome of *Arcobacter* spec.**

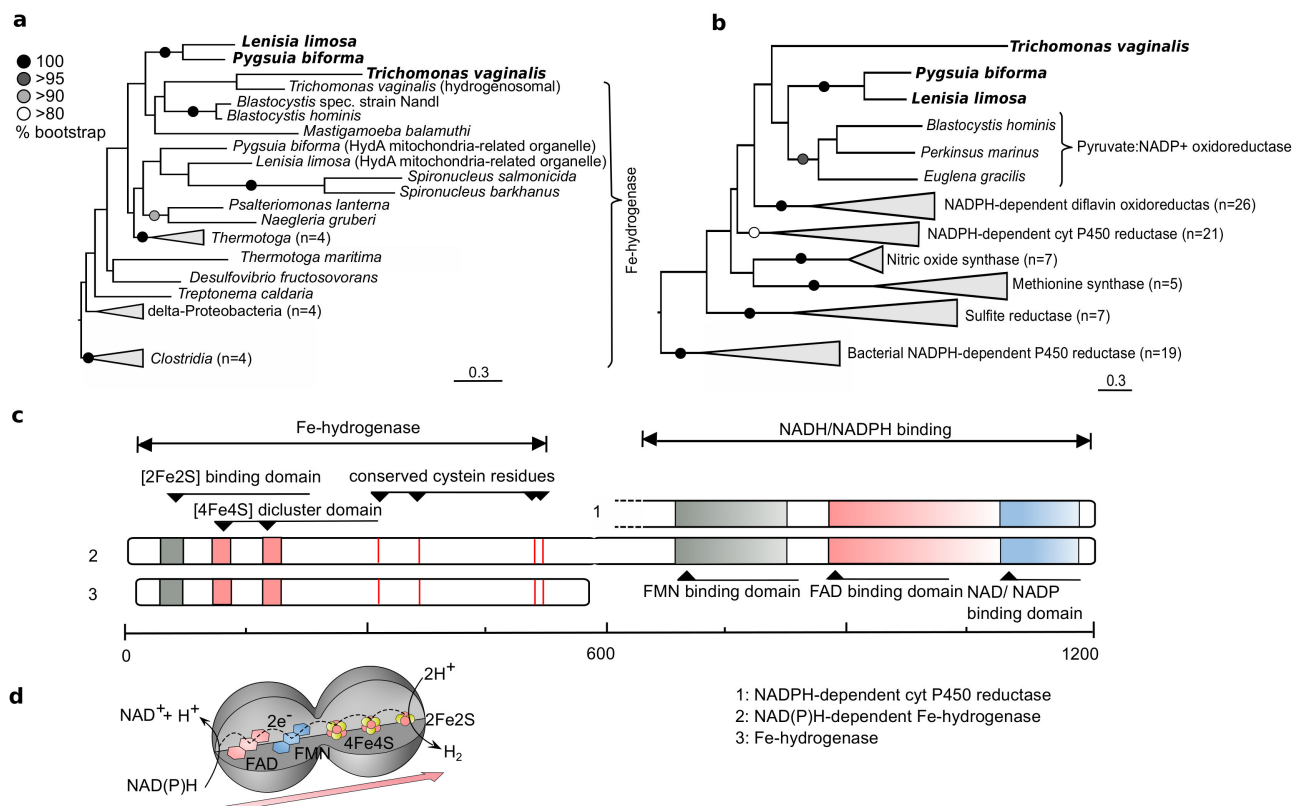
Genome statistics		Check M quality check	
Genome size	3.0 Mb	Marker lineage	Epsilonproteobacteria
%GC	31.94 +/- 2.52	No. of marker genes	447
Average coverage	15.3	Present 1x	428
Contigs	229	Present 2x	18
Shortest contig	1009 bp	Absent	1
Longest contig	528509 bp		
N50	282378	% Completeness	99.6



**The 46.7 Mb genome of *Lenisia limosa***

Genome statistics		Genome quality check	
Genome size	46.7 Mb	Marker lineage	Eukaryota
%GC	41.06 +/- 3.79	No. of marker genes	159
Average coverage	16.3	Present	151
Contigs	11728	Absent	8
Longest contig	135.5 kb		
N50	9095	% Completeness	95

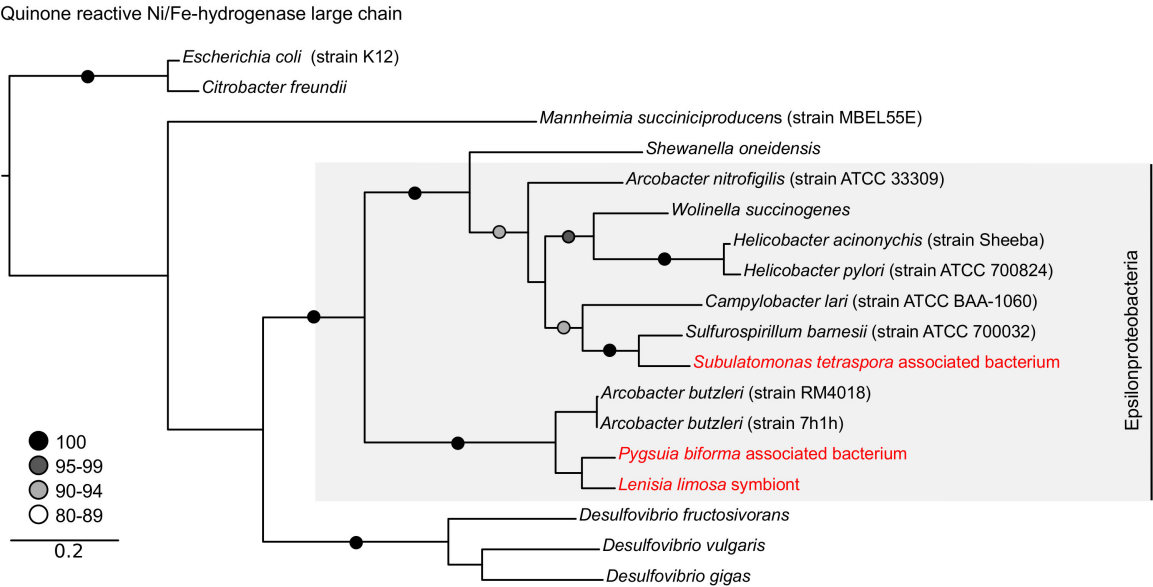
**Extended Data Figure 3 | Genome statistics for *L. limosa* and epibiotic *Arcobacter*.** The pie chart represents the classifications of gene models into functional categories for *Arcobacter*. Gene classifications were performed with the RAST functional annotations and the SEED subsystem database<sup>32</sup>.



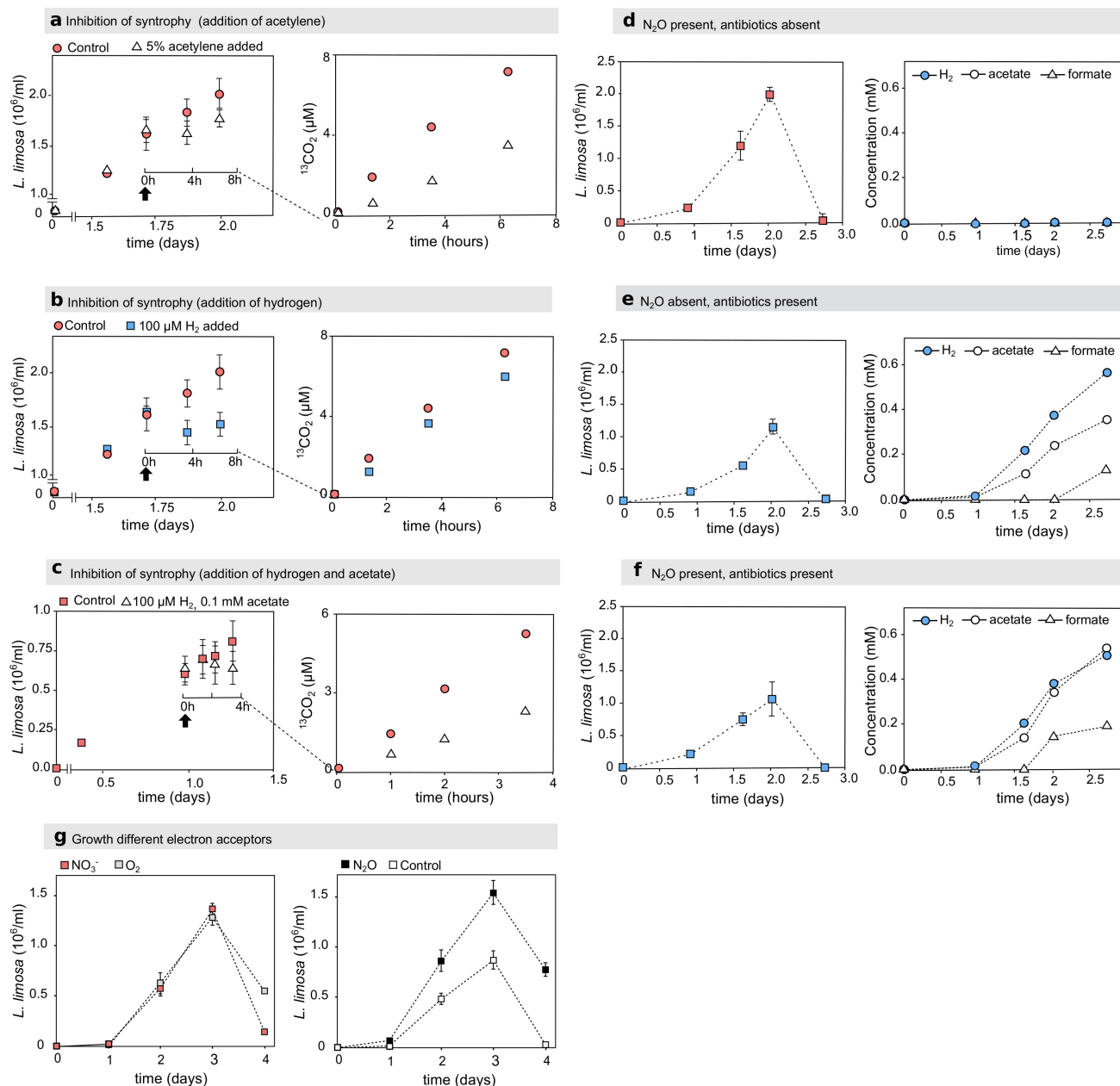
**Extended Data Figure 4 | A new type of NAD(P)H-dependent Fe-hydrogenase.** The genome of *L. limosa* encoded a so far undescribed NAD(P)H-dependent Fe-hydrogenase. Genes with identical domain architecture were also identified in *P. biforma* and *T. vaginalis* (shown in bold type). The scale bars represent substitution rate per site. **a**, Phylogeny of the Fe-hydrogenases domain. **b**, Phylogeny of the NADH/NADP binding domain. Phylogenies were inferred by RAxML using the WAG amino-acid replacement matrix. **c**, Domain architecture of the

NAD(P)H-dependent Fe-hydrogenase (2) compared with the domain architecture of Fe-hydrogenase (3) and the NADPH accepting domain of the cyt P450 reductase (1). The scale bar shows approximate amino-acid positions. **d**, Predicted electron flow within the NAD(P)H-dependent Fe-hydrogenase indicates the capability for a proton-dependent recycling of NAD(P)H. Note: the shape of the model does not intent to depict the actual three-dimensional structure of the protein.



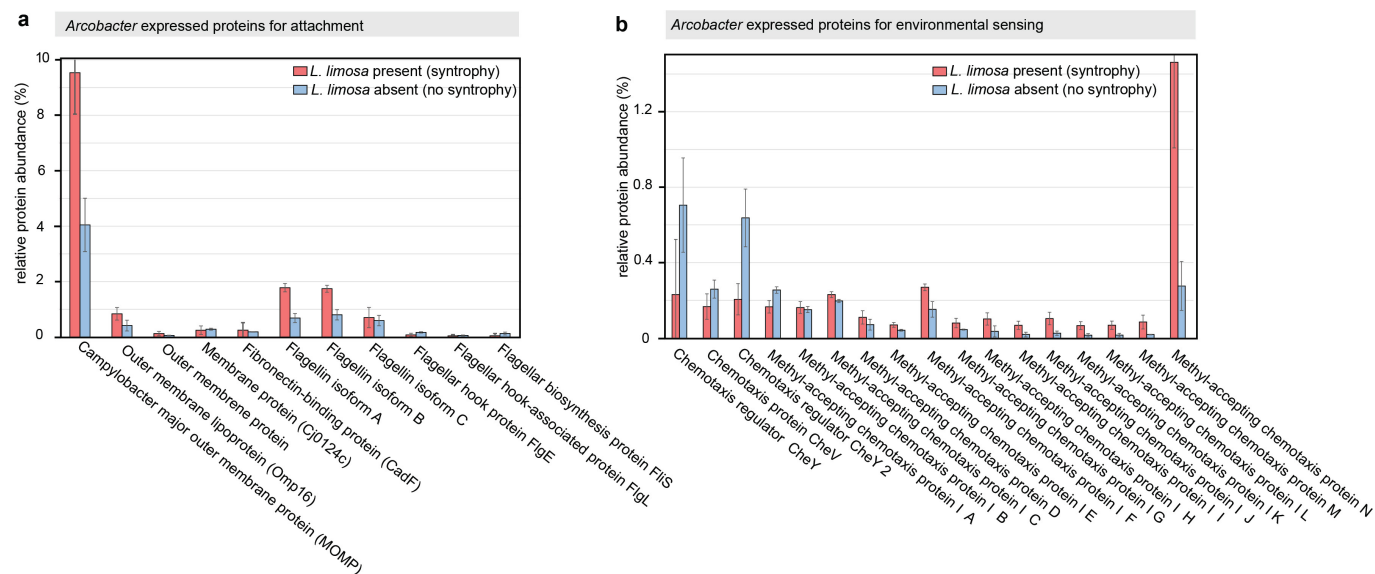


**Extended Data Figure 5 | Maximum likelihood tree of quinone-reactive Ni/Fe-hydrogenases (subunit hydB).** The tree shows the phylogenetic relation of quinone-reactive Ni/Fe-hydrogenases from *Arcobacter* associated with *S. tetraspora*, *P. biforma* and *L. limosa* (indicated in red). Circles represent bootstrap support values for each node. The scale bar represents substitution rate per site.



**Extended Data Figure 6 | The fitness of *L. limosa* depends on its symbiont.** Syntrophy was enabled by the presence of nitrous oxide acting as electron acceptor for bacterial hydrogen oxidation. **a**, Inhibition of nitrous oxide reduction (addition of the competitive inhibitor acetylene, see arrow) led to a reduced growth of *L. limosa* and reduced respiration rates. To monitor respiration rates,  $^{13}\text{C}$ -enriched *Alteromonas* were added together with acetylene. Digestion of  $^{13}\text{C}$ -labelled bacteria by *L. limosa* led to the production of  $^{13}\text{C}$ -bicarbonate, which was measured after conversion to  $^{13}\text{CO}_2$  (right). Similar effects on the growth and respiration rates were observed after adding hydrogen (**b**) or hydrogen and acetate (**c**) to a culture. **d**, Growth of *L. limosa* and production of hydrogen and fatty

acids while growing syntrophically (nitrous oxide present). **e**, Growth of *L. limosa* in the presence of antibacterial antibiotics (nitrous oxide absent). **f**, Growth of *L. limosa* in the presence of antibacterial antibiotics (nitrous oxide present). **g**, Growth of *L. limosa* in the presence of nitrate (2 mM) and oxygen (0.2 mM). Growth of *L. limosa* was compared with a culture that did not contain an electron acceptor for hydrogen oxidation. Each panel shows the results of at least five independent experiments, with cell numbers depicted as averages of seven cell counts per experiment; error bars, s.d.

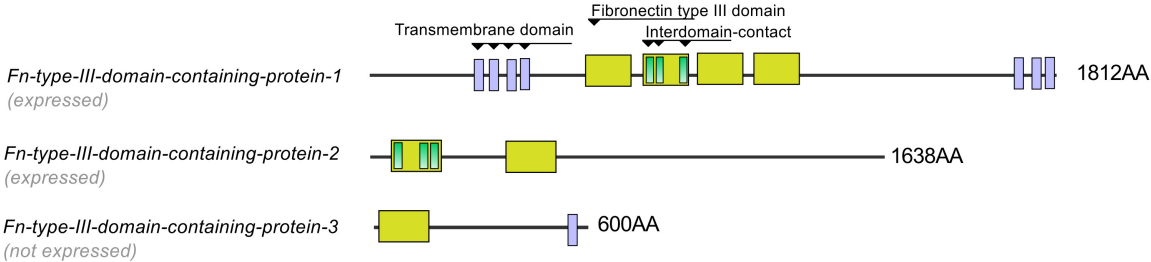


**Extended Data Figure 7 | Expression levels for *Arcobacter* proteins involved in attachment and chemotaxis.** **a.** Expression level of proteins involved in attachment in the presence (red) and absence of *L. limosa* (blue). **b.** Expression level of proteins involved in chemotaxis. Expression

levels were measured and averaged for three independent experiments per treatment (see also Extended Data Fig. 2). Error bars, s.d. See Supplementary Table 1 for gene accession numbers and statistical tests.



Domain architecture fibronectin type III domain containing proteins



**Extended Data Figure 8 | Domain architecture of *L. limosa* fibronectin type III domain-containing proteins.** Protein architectures and conserved protein domains were identified using the SMART protein domain detection tools. See Supplementary Table 1 for gene accession numbers and expression levels.

**Extended Data Table 1 | Potential presence of *Breviatea* and *Breviatea*-associated *Arcobacter* detected in currently available shotgun metagenomes from marine sediments**

				Breviatea (including <i>L. limosa</i> )		Breviatea-associated <i>Arcobacter</i>
SRA accession	F*	Description	#reads	Abundance	P-value	Abundance
				(per 1 million reads)		(per 1 million reads)
SRR2638107	+	sediment, upper 2cm, Barataria Bay (USA)	869558078	0.043	0.0000	0.64
SRR2657558	+	sediment, upper 2cm, Bay Jimmy (USA)	943961688	0.036	0.0000	0.86
SRR2636951	+	sediment, upper 2cm, Barataria Bay (USA)	738401828	0.037	0.0000	0.79
SRR2637699	+	sediment, upper 2cm, Barataria Bay (USA)	986158388	0.028	0.0000	0.13
SRR2657575	+	sediment, upper 2cm, Bay Jimmy (USA)	642092762	0.028	0.0000	3.09
SRR1793930	-	igneous basalt, Louisville Seamounts	65221092	0.123	0.0000	0.02
SRR577221 <sup>†</sup>	+	tidal flat sediment upper 2cm	577090	5.198	0.0000	3.47
SRR577224 <sup>†</sup>	+	tidal flat sediment upper 2cm	682998	4.392	0.0000	10.25
SRR2657076	+	sediment, upper 2cm, Barataria Bay (USA)	164252226	0.055	0.0000	0.15
SRR2657962	+	sediment, upper 2cm, Barataria Bay (USA)	595816470	0.015	0.0006	3.65
SRR577220 <sup>†</sup>	+	tidal flat sediment upper 2cm	581409	1.720	0.0023	6.88
SRR2637322	+	sediment, upper 2cm, Barataria Bay (USA)	144887606	0.028	0.0025	0.22
SRR2656923	+	sediment, upper 2cm, Barataria Bay (USA)	695600554	0.012	0.0052	0.14
SRR2657585	+	sediment, upper 2cm, Barataria Bay (USA)	607701178	0.012	0.0083	0.02
SRR2637706	-	Sediment, 8-12cm, Barataria Bay (USA)	140402750	0.021	0.0164	0.14
SRR2657208	+	sediment, upper 2cm, Terrebonne (USA)	402338618	0.012	0.0173	0.01
SRR2657207	+	sediment, upper 2cm, Barataria Bay (USA)	137181860	0.015	0.0855	0.04
SRR2658004	+	sediment, upper 2cm, Barataria Bay (USA)	162613058	0.012	0.1087	2.64
SRR1627906	-	ocean sediment CostaRica, 32m deep	35468880	0.028	0.1219	0.00
SRR2637708	-	Sediment, 8-12cm, Barataria Bay (USA)	943467314	0.002	0.1668	0.06
SRR2637690	-	Sediment, 8-12cm, Barataria Bay (USA)	596039480	0.005	0.2067	0.22
SRR2658026	-	Sediment, 8-12cm, Barataria Bay (USA)	604747158	0.005	0.2086	0.24
SRR2657594	-	Sediment, 8-12cm, Barataria Bay (USA)	619813316	0.002	0.2113	0.84
SRR2657566	-	Sediment, 8-12cm, Barataria Bay (USA)	607030206	0.002	0.2176	0.09
SRR2657582	+	sediment, upper 2cm, Barataria Bay (USA)	803328968	0.004	0.2229	0.10
SRR2657625	+	sediment, upper 2cm, Barataria Bay (USA)	505336664	0.004	0.2707	4.71
SRR2656927	-	Sediment, 8-12cm, Barataria Bay (USA)	120576078	0.008	0.2960	0.15
SRR2657579	+	sediment, upper 2cm, Barataria Bay (USA)	135244394	0.007	0.3133	0.13
SRR2656924	+	sediment, upper 2cm, Barataria Bay (USA)	143593200	0.007	0.3218	0.10
SRR2656926	+	sediment, upper 2cm, Barataria Bay (USA)	157234396	0.006	0.3339	0.43
SRR2656925	-	Sediment, 8-12cm, Barataria Bay (USA)	618546874	0.000	-	0.01
SRR2657909	-	Sediment, 8-12cm, Barataria Bay (USA)	484269622	0.000	-	0.21
SRR1179191	-	Mahoney Lake (euxinic)	199487463	0.000	-	0.00
SRR2638077	-	Sediment, 8-12cm, Barataria Bay (USA)	196861140	0.000	-	0.03
SRR2657627	+	Sediment, 0-2cm, Barataria Bay (USA)	157919092	0.000	-	2.29
SRR2657590	+	Sediment, 0-2cm, Barataria Bay (USA)	149050036	0.000	-	0.02
SRR2657155	-	Sediment, 8-12cm, Barataria Bay (USA)	121323126	0.000	-	0.17
SRR1627905	-	ocean sediment Costa Rica margin, 2.9m	86658932	0.000	-	0.02
SRR1628696	-	pacific ocean 280m deep, igneous rock	84164942	0.000	-	0.00
SRR1971620	-	Haakon Mosby mud volcano, 3.7m deep	46060794	0.000	-	0.00
SRR1628698	-	pacific ocean 280m deep, igneous rock	45206218	0.000	-	1.26
SRR1628697	-	pacific ocean 280m deep, igneous rock	42072910	0.000	-	0.00
SRR1971621	-	Haakon Mosby mud volcano, 3.7m deep	41183206	0.000	-	0.15
SRR1627907	-	ocean sediment Costa Rica margin, 94m deep	34240066	0.000	-	0.00
SRR1793929	-	pacific ocean, basalt	30470478	0.000	-	0.00
SRR1793931	-	pacific ocean 130m deep, igneous rock	30177380	0.000	-	0.00
SRR1022349	+	marine fish farm sediment	26561056	0.000	-	0.15
SRR1971622	-	Haakon Mosby mud volcano, 2.8m deep	20696408	0.000	-	0.00
SRR1793928	-	pacific ocean, drill fluid	1537452	0.000	-	0.00
SRR577219 <sup>†</sup>	+	tidal flat sediment, upper 5cm	667625	0.000	-	1.50

\*Sediments were grouped into two habitat types: (+) indicates habitats favourable for growth of *Breviatea*, (-) indicates habitats unfavourable for growth of *Breviatea*. The P values (calculated from a binomial distribution with a false positive rate obtained from the unfavourable habitats) are the probabilities that the actual number of reads assigned to *Breviatea* was coincidental and not related to the potential presence of *Breviatea*.

†These metagenomes were obtained from the same site as the inocula for the enrichment of *L. limosa* in the present study.

# Universality of human microbial dynamics

Amir Bashan<sup>1</sup>, Travis E. Gibson<sup>1</sup>, Jonathan Friedman<sup>2</sup>, Vincent J. Carey<sup>1</sup>, Scott T. Weiss<sup>1</sup>, Elizabeth L. Hohmann<sup>3</sup> & Yang-Yu Liu<sup>1,4</sup>

Human-associated microbial communities have a crucial role in determining our health and well-being<sup>1,2</sup>, and this has led to the continuing development of microbiome-based therapies<sup>3</sup> such as faecal microbiota transplantation<sup>4,5</sup>. These microbial communities are very complex, dynamic<sup>6</sup> and highly personalized ecosystems<sup>3,7</sup>, exhibiting a high degree of inter-individual variability in both species assemblages<sup>8</sup> and abundance profiles<sup>9</sup>. It is not known whether the underlying ecological dynamics of these communities, which can be parameterized by growth rates, and intra- and inter-species interactions in population dynamics models<sup>10</sup>, are largely host-independent (that is, universal) or host-specific. If the inter-individual variability reflects host-specific dynamics due to differences in host lifestyle<sup>11</sup>, physiology<sup>12</sup> or genetics<sup>13</sup>, then generic microbiome manipulations may have unintended consequences, rendering them ineffective or even detrimental. Alternatively, microbial ecosystems of different subjects may exhibit universal dynamics, with the inter-individual variability mainly originating from differences in the sets of colonizing species<sup>7,14</sup>. Here we develop a new computational method to characterize human microbial dynamics. By applying this method to cross-sectional data from two large-scale metagenomic studies—the Human Microbiome Project<sup>9,15</sup> and the Student Microbiome Project<sup>16</sup>—we show that gut and mouth microbiomes display pronounced universal dynamics, whereas communities associated with certain skin sites are probably shaped by differences in the host environment. Notably, the universality of gut microbial dynamics is not observed in subjects with recurrent *Clostridium difficile* infection<sup>17</sup> but is observed in the same set of subjects after faecal microbiota transplantation. These results fundamentally improve our understanding of the processes that shape human microbial ecosystems, and pave the way to designing general microbiome-based therapies<sup>18</sup>.

The underlying dynamics of a microbial ecosystem, that is, the ecological interactions that govern its change, equilibrium and stability, can be represented by a population dynamic model

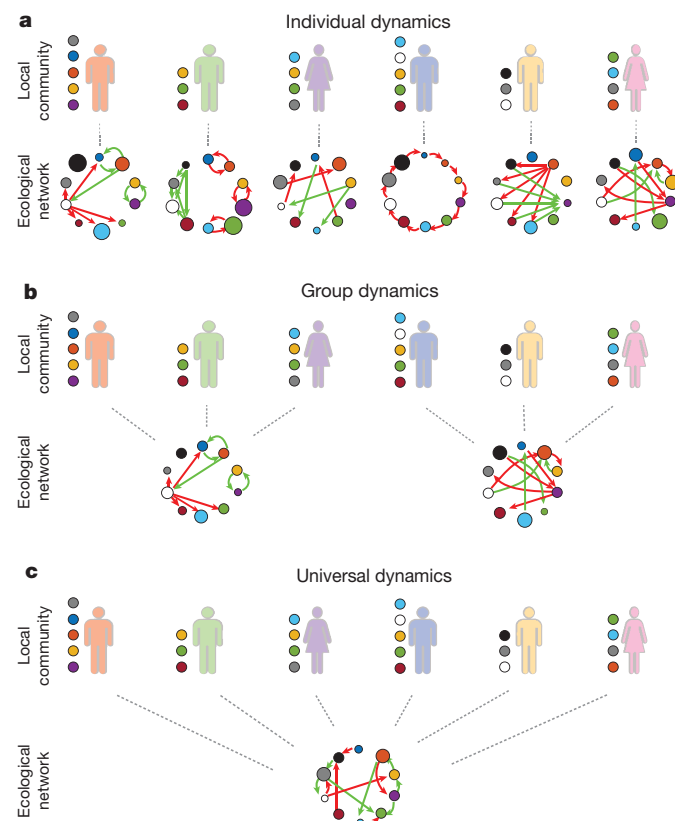
$$\dot{\mathbf{x}}^{(\nu)} = \mathbf{f}(\mathbf{x}^{(\nu)}; \boldsymbol{\Theta}^{(\nu)}) \quad (1)$$

which describes the time-dependent abundance profile  $\mathbf{x}^{(\nu)}(t) = (x_1^{(\nu)}(t), \dots, x_N^{(\nu)}(t))$  of  $N$  microbial species present in a particular body site of subject  $\nu$ . Here,  $\mathbf{f}(\mathbf{x}^{(\nu)}; \boldsymbol{\Theta}^{(\nu)})$  is typically a nonlinear function and  $\boldsymbol{\Theta}^{(\nu)}$  captures all the ecological parameters, that is, growth rates, and intra- and inter-species interactions. Those parameters may generally depend on host-independent factors, such as biochemical processes and microbial metabolic pathways<sup>19</sup>; and on host-specific ones, such as nutrient intake<sup>20</sup> and host genetic make-up<sup>13</sup>.

Three fundamental cases could represent the dynamics of  $M$  healthy subjects: (1) individual dynamics, in which the ecological parameters are different in different subjects, that is,  $\boldsymbol{\Theta}^{(1)} \neq \dots \neq \boldsymbol{\Theta}^{(M)}$ ; (2) group dynamics, in which subjects can be classified into  $K$  groups ( $K \ll M$ ) on the basis of certain host factors and subjects in the same group share the same set of parameters, that is,  $\boldsymbol{\Theta}^{(\nu)} = \boldsymbol{\Theta}^P$  for all subjects in

group  $P$  ( $P = 1, \dots, K$ ); and (3) universal dynamics, in which all the subjects share the same set of parameters, that is,  $\boldsymbol{\Theta}^{(\nu)} = \boldsymbol{\Theta}$  for all subjects. If we represent the ecological parameters, such as the inter-species interactions, in a directed, weighted ecological network, the above three cases can be easily visualized (see Fig. 1).

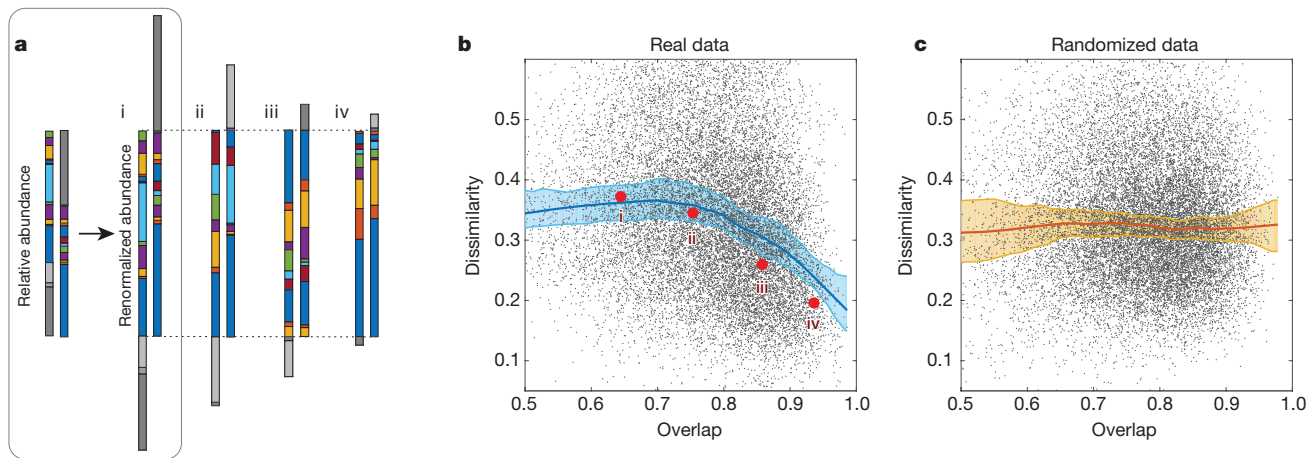
Despite its crucial consequences, we do not know which case best represents the microbial ecosystems of healthy individuals. Addressing this question is vital for developing microbiome-based therapies<sup>3,18</sup>. Indeed, if the dynamics are universal, the inter-personal variability stems solely from the different assemblages of colonizing species in different individuals. We can then design



**Figure 1 | Alternative scenarios of microbial dynamics across different healthy subjects.** Microbial dynamics captured by equation (1) is simply represented by an ecological network, in which nodes represent species (with node sizes proportional to growth rates) and edges represent inter-species interactions (with green and red arrows representing excitatory and inhibitory interactions, respectively). Different subjects typically have different species assemblages, represented by coloured circles near each subject. **a**, The underlying dynamics/network is unique for each subject. **b**, Subjects within the same group share the same dynamics/network that is significantly different from that of other groups. **c**, Different subjects have the same underlying dynamics/network.

<sup>1</sup>Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>2</sup>Physics of Living Systems, Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. <sup>3</sup>Infectious Disease Division, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>4</sup>Center for Cancer Systems Biology, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA.





**Figure 2 | Higher overlap of microbial communities is associated with lower dissimilarity.** **a**, Four gut microbial sample pairs (i–iv) represented by stacked bars at the genus level. For each sample pair, their shared genera are coloured while non-shared genera are shown in grey. **b**, DOC (in dark blue) of gut microbial sample pairs from the HMP study

general interventions to control the microbial state (in terms of species assemblage and abundance profile) of different individuals. By contrast, if the dynamics are strongly host-specific, we must design truly personalized interventions, which need to consider not only the unique microbial state of an individual but also the unique dynamics of the underlying microbial ecosystem. In addition, host-specific dynamics, if they exist, raise a major safety concern for faecal microbiota transplantation (FMT), because although the healthy microbiota are stable in the donor's gut, they may be shifted to an undesired state in the recipient's gut.

The ideal approach to addressing this fundamental question would be to infer the dynamic model captured by equation (1) for a large number of healthy individuals from temporal metagenomic data, and then compare the system parameters  $\Theta^{(\nu)}$  directly. However, empirical parameterization of the exact functional form of  $f(\mathbf{x}^{(\nu)}; \Theta^{(\nu)})$  is extremely difficult for complex ecological systems. Furthermore, inferring the system parameters typically requires high-quality time series data and well-designed experiments to ensure the system parameters are identifiable<sup>21</sup>. Such data sets are not currently available. A conventional correlation analysis of cross-sectional data cannot address this question either, because it only captures effective (or indirect) interactions and is subject to spurious correlations due to the compositionality of relative abundances in genomic survey data<sup>22</sup>.

To overcome these issues, we developed a novel method to detect 'fingerprints' of universal microbial dynamics. This is achieved by restricting ourselves to answer the question of whether the dynamics are universal or not, rather than the broader and harder question of what the dynamics are. The key idea is that when comparing microbial communities (samples) from different subjects, we distinguish between two contributors to the inter-individual variability: the difference in species assemblages and the difference in abundance profiles. We quantify those two contributors by:  $O(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ , the overlap of the species assemblages, calculated from the relative abundances of the shared species; and  $D(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ , the dissimilarity between the renormalized abundance profiles of the shared species (see Methods). Note that the two measures (overlap and dissimilarity) are not *a priori* dependent on each other. Indeed,  $D(\hat{\mathbf{x}}, \hat{\mathbf{y}})$  is mathematically not constrained by any value of  $O(\hat{\mathbf{x}}, \hat{\mathbf{y}}) > 0$  (see Supplementary Information section 1.2.1 for details). Hence any constraints of  $D(\hat{\mathbf{x}}, \hat{\mathbf{y}})$  by  $O(\hat{\mathbf{x}}, \hat{\mathbf{y}})$  observed from real data deserve our attention and may have ecological interpretations (see Fig. 2a, b).

To compare samples systematically from a given microbiome data set, we first calculate the overlap and dissimilarity of all the sample pairs and represent each sample pair as a point in the

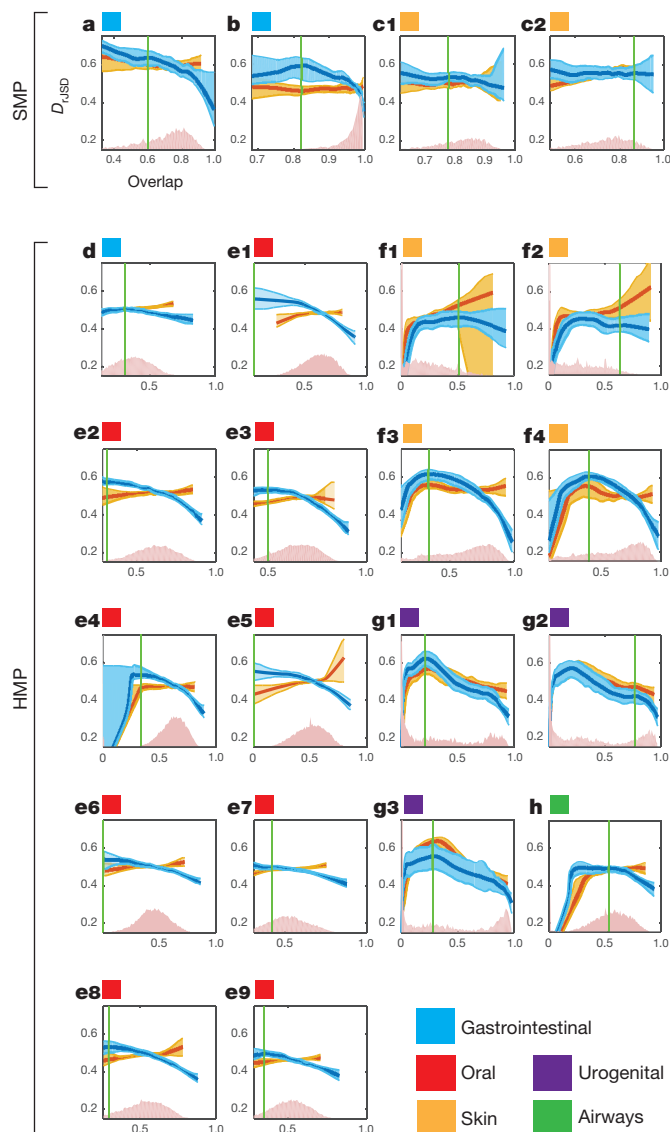
( $M = 190$  samples). Grey dots represent all the 17,955 sample pairs. **c**, DOC (in dark red) of the randomized samples is flat. In **b** and **c**, and throughout the paper, shaded area indicates the range of the 94% confidence interval (see Methods).

dissimilarity–overlap plane. We then perform nonparametric regression and bootstrap sampling to calculate the dissimilarity–overlap curve (DOC) and its confidence interval (see Fig. 2b and Methods). In the case of (i) individual dynamics, or (ii) universal dynamics but without inter-species interactions, a flat DOC is expected (see Supplementary Information section 1.2.3). By contrast, for systems with universal dynamics and inter-species interactions, we expect the corresponding DOC to display a characteristic feature: a negative slope in the high-overlap region; that is, abundance profiles of sample pairs become more similar as their overlap becomes higher (see Supplementary Information section 1.2 and Extended Data Fig. 1). A negative slope can also be seen in the DOC of microbial communities characterized by group dynamics. The existence of such group dynamics however can be easily detected by standard ordination techniques and clustering analysis<sup>23,24</sup> and hence ruled out (see Extended Data Fig. 2). Note that the DOC analysis described above is not affected by the compositionality of the genomic survey data and requires neither time series data nor any *a priori* knowledge of the specific ecological dynamics. Instead, it only relies on a few reasonable assumptions (see Methods).

To verify our DOC analysis, we first applied it to synthetic data generated from the canonical generalized Lotka–Volterra (GLV) model, which has been used for predictive modelling of the intestinal microbiota<sup>25–27</sup>. Extended Data Figure 3 shows that in the case of universal dynamics with strong inter-species interactions, the DOC displays a clear negative slope in the high-overlap region. By contrast, in the case of individual dynamics or universal dynamics without inter-species interactions, a flat DOC is observed.

To verify the DOC analysis directly using real data, we analysed longitudinal gut microbial samples of four healthy individuals from two microbiome studies<sup>11,28</sup>. For each individual, we expect a highly universal microbial dynamics throughout the period of measurement; that is, the ecological parameters  $\Theta^{(\nu)}$  of the corresponding microbial community are largely time-invariant. We found that the DOCs of all four subjects show a clear negative slope in the high-overlap region (Extended Data Fig. 4), consistent with our expectation.

Next, we systematically analysed cross-sectional microbial samples of different body sites from two large-scale metagenomic studies, the Human Microbiome Project (HMP)<sup>9,15</sup> and the Student Microbiome Project (SMP)<sup>16</sup>. The results are shown in Fig. 3 and Extended Data Fig. 5. In Fig. 3, for each body site the DOCs calculated from real and randomized samples are shown in dark blue and red, respectively. The overlap distributions of the real between-subjects sample pairs are shown in pink. Note that the characteristic overlap in a particular body



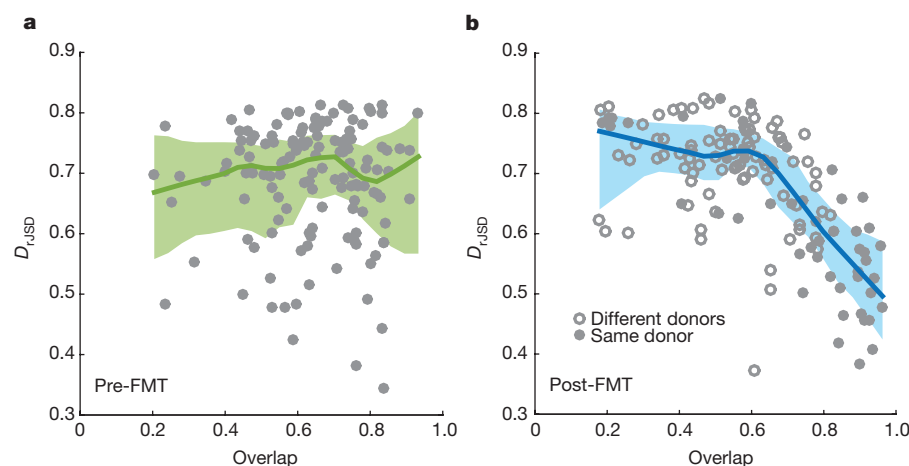
site is different in the two studies. For example, the average overlap between HMP gut samples is about 0.4 and between SMP samples is about 0.75. To account for this fact and to compare the DOCs fairly across different body sites and studies, we used two different measures to quantify the universality (see Methods). Notably, although these two measures quantify different features of the DOC analysis, the body-sites stratification pattern is consistent across the two measures and

**Figure 3 | Detecting universality of microbial dynamics in different body sites.** a–h, We calculated DOCs for real (dark blue) and randomized (dark red) samples of two data sets: (1) SMP: gut (a), tongue (b), forehead skin (c1), palm skin (c2); (2) HMP: gut (d), tongue dorsum (e1), attached keratinized gingiva (e2), buccal mucosa (e3), hard palate (e4), palatine tonsils (e5), subgingival plaque (e6), supragingival plaque (e7), throat (e8), saliva (e9), left/right antecubital fossa (f1/f2), left/right retroauricular crease (f3/f4), vaginal introitus (g1), mid-vagina (g2), posterior fornix (g3), anterior nares (h). The overlap distributions of the real between-subjects sample pairs are shown in pink. The vertical green line represents the change point (see Methods).

the two studied data sets (Extended Data Fig. 6). In particular, the negative slope of DOC is most significantly observed in samples from the gut and mouth and least observed in samples from hand skin (palm and elbow). These findings strongly suggest the existence of universal dynamics characterized by inter-species interactions in the gut and mouth microbiomes.

An alternative explanation for the observed negative slope of the DOC for gut and mouth microbiomes of healthy subjects could be that some host factors not only select for the presence of certain microbes but also drive their relative abundances by enforcing certain optimally adapted compositions. To test this alternative explanation, we systematically analysed microbial samples while controlling for the effect of several leading candidates for potential confounding factors, for example, body mass index, age, long-term dietary pattern and stool consistency. We found that as long as their values are in the normal range those factors cannot explain the observed DOC pattern (see Extended Data Figs 7 and 8). Hence, the alternative explanation for the negative slope in DOC is unlikely to be true. Of course, with currently available data sets we cannot possibly account for all other confounders, such as drugs, genetics, inflammation, or their combinations. More data sets will be needed to test this intriguing alternative explanation.

The above results of healthy subjects raise an interesting question: does the universality of microbial dynamics also exist in subjects with disrupted microbiomes? To address this question, we applied the DOC analysis to microbial samples of 17 subjects with recurrent *Clostridium difficile* infection (rCDI) and the same set of subjects after FMT<sup>17</sup>. *Clostridium difficile* is an opportunistic pathogen that causes disease worldwide and greatly increases morbidity and mortality in hospitalized patients. Fortunately, FMT is very efficacious in treating patients with rCDI, with pronounced clinical improvement even after a single treatment<sup>29</sup>. We found that the dissimilarity between rCDI subjects is largely independent of their species overlap, rendering a flat DOC (Fig. 4a). By contrast, after FMT (median, 4 days) the DOC displays a pronounced negative slope in the high-overlap region (Fig. 4b), suggesting a universal gut microbial dynamics. FMT treatments show the flexibility of microbial communities and their adaptation to composition changes. Our result suggests that this adaptive behaviour may be associated with the observed universal microbial dynamics after FMT.



**Figure 4 | DOC analysis of human subjects with rCDI.** a, Before FMT, the DOC (dark green line) of the rCDI subjects is nearly flat. b, After FMT, the DOC (dark blue line) displays a pronounced negative slope in the high-overlap region. We denoted a subject pair as a solid (or hollow) circle if the two subjects received FMT from the same donor (or two different donors). Notably, solid circles spread over a wide range of overlap values, suggesting that even if two subjects share the same donor, their post-FMT microbiomes may still display strong inter-individual variability.

Finally, we anticipate that applying our DOC analysis to subjects with other diseases (especially non-gastrointestinal diseases) or infants at different developmental stages will offer deeper insights into how dynamical processes shape human microbial ecosystems. The developed DOC analysis can also be directly applied to other microbial ecosystems—for example, the microbiome of soil, ocean, lakes, phyllosphere/rhizosphere and fermenters—to detect the universality of the underlying ecological dynamics (see Extended Data Fig. 9). This sheds light on the design of more advanced methods to extract dynamical information from microbial data.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 24 August 2015; accepted 9 May 2016.**

1. Cho, I. & Blaser, M. J. The human microbiome: at the interface of health and disease. *Nat. Rev. Genet.* **13**, 260–270 (2012).
2. Pflughoeft, K. J. & Versalovic, J. Human microbiome in health and disease. *Annu. Rev. Pathol.* **7**, 99–122 (2012).
3. Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K. & Knight, R. Diversity, stability and resilience of the human gut microbiota. *Nature* **489**, 220–230 (2012).
4. Borody, T. J. & Khoruts, A. Fecal microbiota transplantation and emerging applications. *Nat. Rev. Gastroenterol. Hepatol.* **9**, 88–96 (2011).
5. Aroniadis, O. C. & Brandt, L. J. Fecal microbiota transplantation: past, present and future. *Curr. Opin. Gastroenterol.* **29**, 79–84 (2013).
6. Gerber, G. K. The dynamic microbiome. *FEBS Lett.* **588**, 4131–4139 (2014).
7. Costello, E. K., Stagaman, K., Dethlefsen, L., Bohannan, B. J. M. & Relman, D. a. The application of ecological theory toward an understanding of the human microbiome. *Science* **336**, 1255–1262 (2012).
8. Franzosa, E. A. *et al.* Identifying personal microbiomes using metagenomic codes. *Proc. Natl Acad. Sci. USA* **112**, E2930–E2938 (2015).
9. The Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012).
10. Bucci, V. & Xavier, J. B. Towards predictive models of the human gut microbiome. *J. Mol. Biol.* **426**, 3907–3916 (2014).
11. David, L. A. *et al.* Host lifestyle affects human microbiota on daily timescales. *Genome Biol.* **15**, R89 (2014).
12. Sommer, F. & Backhed, F. The gut microbiota—masters of host development and physiology. *Nat. Rev. Microbiol.* **11**, 227–238 (2013).
13. Goodrich, J. K. *et al.* Human genetics shape the gut microbiome. *Cell* **159**, 789–799 (2014).
14. Walter, J. & Ley, R. The human gut microbiome: ecology and recent evolutionary changes. *Annu. Rev. Microbiol.* **65**, 411–429 (2011).
15. The Human Microbiome Project Consortium. A framework for human microbiome research. *Nature* **486**, 215–221 (2012).
16. Flores, G. E. *et al.* Temporal variability is a personalized feature of the human microbiome. *Genome Biol.* **15**, 531 (2014).
17. Youngster, I. *et al.* Fecal microbiota transplant for relapsing *Clostridium difficile* infection using a frozen inoculum from unrelated donors: a randomized, open-label, controlled pilot study. *Clin. Infect. Dis.* **58**, 1515–1522 (2014).

18. Lemon, K. P., Armitage, G. C., Relman, D. a. & Fischbach, M. a. Microbiota-targeted therapies: an ecological perspective. *Sci. Transl. Med.* **4**, 137rv135 (2012).
19. Levy, R. & Borenstein, E. Metabolic modeling of species interaction in the human microbiome elucidates community-level assembly rules. *Proc. Natl Acad. Sci. USA* **110**, 12804–12809 (2013).
20. Jumpertz, R. *et al.* Energy-balance studies reveal associations between gut microbes, caloric load, and nutrient absorption in humans. *Am. J. Clin. Nutr.* **94**, 58–65 (2011).
21. Faust, K. & Raes, J. Microbial interactions: from networks to models. *Nat. Rev. Microbiol.* **10**, 538–550 (2012).
22. Friedman, J. & Alm, E. J. Inferring correlation networks from genomic survey data. *PLoS Comput. Biol.* **8**, e1002687 (2012).
23. Koren, O. *et al.* A guide to enterotypes across the human body: meta-analysis of microbial community structures in human microbiome datasets. *PLoS Comput. Biol.* **9**, e1002863 (2013).
24. Gibson, T. E., Bashan, A., Cao, H.-T., Weiss, S. T. & Liu, Y.-Y. On the origins and control of community types in the human microbiome. *PLoS Comput. Biol.* **12**, e1004688 (2016).
25. Stein, R. R. *et al.* Ecological modeling from time-series inference: insight into dynamics and stability of intestinal microbiota. *PLoS Comput. Biol.* **9**, e1003388 (2013).
26. Fisher, C. K. & Mehta, P. Identifying keystone species in the human gut microbiome from metagenomic timeseries using sparse linear regression. *PLoS ONE* **9**, e102451 (2014).
27. Buffie, C. G. *et al.* Precision microbiome reconstitution restores bile acid mediated resistance to *Clostridium difficile*. *Nature* **517**, 205–208 (2015).
28. Caporaso, J. G. *et al.* Moving pictures of the human microbiome. *Genome Biol.* **12**, R50 (2011).
29. Kassam, Z., Lee, C. H., Yuan, Y. & Hunt, R. H. Fecal microbiota transplantation for *Clostridium difficile* infection: systematic review and meta-analysis. *Am. J. Gastroenterol.* **108**, 500–508 (2013).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank E. K. Silverman, G. Weinstock, C. Huttenhower, R. Knight, G. Ackermann, D. Del Vecchio, D. Lauffenburger, G. Abu-Ali, J. Sordillo, M. McGeachie, and J. Gore for discussions. Special thanks to A.-L. Barabási and J. Loscalzo for careful reading of the manuscript. This work was partially supported by the John Templeton Foundation (award number 51977) and National Institutes of Health (R01 HL091528).

**Author Contributions** Y.-Y.L. conceived and designed the project. A.B. developed the DOC analysis, performed numerical simulations, and analysed all the real data. A.B. and Y.-Y.L. performed analytical calculations. A.B. and V.J.C. performed statistical tests. All authors analysed the results. A.B. and Y.-Y.L. wrote the manuscript. All authors edited the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Y.-Y.L. ([yyyl@channing.harvard.edu](mailto:yyyl@channing.harvard.edu)).

**Reviewer Information** Nature thanks F. He, P. Rohani and the other anonymous reviewer(s) for their contribution to the peer review of this work.



## METHODS

All the data sets analysed in this work have been published (see Methods section 'Human microbiome data sets analysed in this work' for details). The original experiments and corresponding power analysis have been reported in previous publications.

**Overlap between species assemblages.** Consider two microbial samples, represented by two abundance vectors  $\mathbf{x} = (x_1, \dots, x_N) \in \mathbb{R}^N$  and  $\mathbf{y} = (y_1, \dots, y_N) \in \mathbb{R}^N$ . For genomic survey data of the human microbiome, only the relative abundances are known. Hence, we are dealing with the relative abundance profiles  $\tilde{\mathbf{x}} = (\tilde{x}_1, \dots, \tilde{x}_N)$  and  $\tilde{\mathbf{y}} = (\tilde{y}_1, \dots, \tilde{y}_N)$ , where  $\tilde{x}_i \equiv \frac{x_i}{\sum_{j=1}^N x_j}$  and  $\tilde{y}_i \equiv \frac{y_i}{\sum_{j=1}^N y_j}$ . To quantify the similarity

of the species assemblages (sets) of the two samples, denoted as  $X = \{i | x_i > 0\}$  and  $Y = \{i | y_i > 0\}$ , we defined the overlap measure  $O(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) \equiv \sum_{i \in S} \frac{\tilde{x}_i + \tilde{y}_i}{2}$ , where

$S \equiv X \cap Y$  is the set of shared species present in both samples. In case  $S$  is empty,  $O(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = 0$ . If  $S = \{1, \dots, N\}$ , that is, all the species in  $X$  and  $Y$  are shared, then  $O(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = 1$ , but the abundance profiles  $\tilde{\mathbf{x}}$  and  $\tilde{\mathbf{y}}$  can still be very different. In the extreme case when the relative abundance is the same for all species in  $X$  and  $Y$ , the overlap measure can be written as a function of the classical Jaccard index. Yet, there are many advantages of using the overlap measure, instead of the Jaccard index, in our analysis (see Supplementary Information section 1.1.4).

**Dissimilarity between abundance profiles.** To compare the abundance profiles of two samples, we first renormalize the relative abundances of only the shared species (in set  $S$ ), yielding  $\hat{\mathbf{x}} = \{\hat{x}_i\}_{i \in S}$  and  $\hat{\mathbf{y}} = \{\hat{y}_i\}_{i \in S}$ . Here

$$\hat{x}_i \equiv \frac{\tilde{x}_i}{\sum_{j \in S} \tilde{x}_j} = \frac{x_i / \sum_{k \in X} x_k}{\sum_{j \in S} x_j / \sum_{k \in X} x_k} = \frac{x_i}{\sum_{j \in S} x_j} \quad \text{and} \quad \hat{y}_i \text{ is defined similarly. This way we}$$

remove the spurious dependence between the relative abundances of the shared and the non-shared species. More importantly, this renormalization assures that the calculated dissimilarity measure is mathematically independent of the overlap measure (see Supplementary Information section 1.2.1). The dissimilarity is then evaluated via the root Jensen–Shannon divergence (rJSD) measure

$$D(\hat{\mathbf{x}}, \hat{\mathbf{y}}) = D_{\text{rJSD}}(\hat{\mathbf{x}}, \hat{\mathbf{y}}) \equiv \left[ \frac{D_{\text{KL}}(\hat{\mathbf{x}}, \mathbf{m}) + D_{\text{KL}}(\hat{\mathbf{y}}, \mathbf{m})}{2} \right]^{\frac{1}{2}}$$

in which  $\mathbf{m} \equiv \frac{\hat{\mathbf{x}} + \hat{\mathbf{y}}}{2}$  and  $D_{\text{KL}}(\hat{\mathbf{x}}, \hat{\mathbf{y}}) \equiv \sum_{i \in S} \hat{x}_i \log \frac{\hat{x}_i}{\hat{y}_i}$  is the Kullback–Leibler divergence between  $\hat{\mathbf{x}}$  and  $\hat{\mathbf{y}}$ . The dissimilarity can also be evaluated via any other classical dissimilarity measures in ecology and biology, for example, Bray–Curtis dissimilarity, Yue–Clayton dissimilarity, and the negative Spearman correlation (see Extended Data Fig. 5). In this work, we focused on rJSD because it is a distance metric that satisfies non-negativity, identity, symmetry and triangle inequality (see Supplementary Information section 1.1.1). Comparing sample pairs on the basis of phylogenetic information, for example, using weighted- and unweighted-UniFrac<sup>30</sup> as quantitative and qualitative measures, respectively, has the potential to provide better insight on the communities' dissimilarity–overlap behaviour. However, since the weighted- and unweighted-UniFrac are not mathematically independent, they cannot be trivially integrated into our DOC analysis.

**DOC.** To compare sample pairs systematically with a wide range of overlap values and analyse their dissimilarity–overlap relations, we calculate the overlap and dissimilarity of all the sample pairs from a given set of microbiome samples and represent each sample pair as a point in the dissimilarity–overlap plane. We then use the robust LOWESS (locally weighted scatterplot smoothing) method, a standard non-parametric regression method that is resistant to outliers, to calculate the DOC.

To get the confidence interval, we use the following bootstrap technique.

(1) From a data set of  $M$  samples we calculate the overlap and dissimilarity of the  $M(M-1)/2$  sample pairs, represented as  $M(M-1)/2$  points in the overlap–dissimilarity plane. (2) In each bootstrap realization, we resample a new set  $K = \{k_1, \dots, k_M\}$  from the  $M$  original samples with replacement. Some of the original samples might not be included and some could be sampled more than once. (3) We create a new cloud of points  $C$ : a point associated with sample pair  $(i, j)$  is included in  $C$  only if both  $i, j \in K$ , while a point is chosen several times if the sample  $i$  or  $j$  were resampled more than once in  $K$ . (4) A new DOC is calculated for  $C$  using the robust LOWESS method. We set the smoothing parameter ('span') to be 0.2. (5) We repeat steps (2)–(4)  $T$  times to create  $T$  DOCs. (6) The 3rd and 97th percentiles of the  $T$  curves represent the 94% confidence interval for the DOC. In this work, we chose  $T = 100$ .

**Assumptions of the DOC analysis.** There are two reasonable assumptions underlying the DOC analysis. First, the abundance profiles of the samples should represent the steady states of the microbial ecosystem and hence the fixed points of the underlying dynamics that satisfy  $\dot{\mathbf{x}} = \mathbf{0}$ . This assumption is fairly reasonable because human gut microbiota is a relatively resilient ecosystem<sup>3</sup>, and until the next large perturbation (for example, antibiotic administration or dramatic dietary

change) is introduced, the system remains stable for months and possibly even years<sup>11,28,31</sup>. Second, if two communities have the same species assemblages and the same abundance profile (steady state), then the two communities have the same microbial dynamics. Mathematically, this is not necessarily true, because different dynamical systems can give rise to an identical steady state or fixed point. Yet, given the large number of species and all the other levels of complexity in their interactions, the possibility of having different dynamics with the same fixed point is very unlikely. Indeed, universal dynamics is the most plausible explanation for the observed pattern, that is, the negative slope of DOC in the high-overlap region.

**Limitations of the DOC analysis.** We point out that for overlap values close to zero, a positive slope may occur as the artefact of dissimilarity between relative abundance profiles with small number of species (see Fig. 3e4, f1–f4, g1–3, h, Extended Data Fig. 10, and Supplementary Information section 1.1.3).

We also emphasize that a flat DOC does not completely rule out the possibility of universal dynamics. For example, the DOC of the gut microbiome samples of rCDI patients is flat (Fig. 4a). There are two possibilities. First, the universality of microbial dynamics found in healthy subjects (Fig. 3a, d) is completely lost in the rCDI subjects, owing to the infection and/or the dysbiosis caused by the extensive inciting antibiotic treatment. Second, the possibly universal microbial dynamics of the rCDI subjects are just undetectable by the DOC analysis. This could be due to the extremely liquid stool samples of the rCDI subjects that suffer from diarrhoea, as stool consistency has been found to be strongly correlated with the gut microbiota compositions<sup>32,33</sup>. It is also possible that the abundance profiles of rCDI subjects are markedly varying over time and hence do not represent the steady states of the underlying microbial ecosystem (although a mouse infection model does not seem to support this hypothesis<sup>34</sup>).

If true multi-stability exists, that is, multiple stable states (abundance profiles) are associated with the same set of species present in the same environment, then our DOC analysis may not detect it. However, true multi-stability in human-associated microbial communities has not been demonstrated experimentally, partially because any subtle differences in the species assemblages can drive those microbial communities<sup>35</sup>.

In summary, our DOC analysis detects universal dynamics under certain conditions. More precisely, it provides a means of discriminating dynamics into universal or possibly universal.

**Universality measures and statistical test.** Note that the DOCs of different data sets/studies must be compared with caution, especially if the microbial samples were preprocessed by different pipelines<sup>36</sup>, for example, with different operational taxonomic unit (OTU) clustering thresholds, or different OTU picking methods. As shown in Fig. 3, the characteristic overlap in a particular body site is different in different studies. For example, the average overlap between HMP gut samples is about 0.4 and between SMP samples is about 0.75. To account for this fact and to compare the DOCs fairly across different body sites and studies, we used two different measures to quantify the universality.

(1)  $f_{\text{ns}}$ . For each cohort we determined the fraction of data points for which the DOC displays a negative slope, denoted as  $f_{\text{ns}}$ . Specifically, for a given DOC calculated from a cohort of  $M$  microbial samples, we first detected the 'change point'  $O_c$  such that  $\frac{dy(O)}{dO} < 0$  for any  $O > O_c$ , in which  $y(O)$  is a smoothed curve of the DOC (for example, using the default 'smooth' function of Matlab with moving average over 5 neighbours). Then,  $f_{\text{ns}}$  is defined as  $f_{\text{ns}} \equiv \frac{\text{number of sample pairs with } O > O_c}{\text{total number of sample pairs}}$ . In Fig. 3a–h, this is the area of the overlap distribution to the right of the green vertical line (the change point  $O_c$ ). The results of  $f_{\text{ns}}$  for different body sites are shown in Extended Data Fig. 6a.

(2)  $P$  value. To estimate the slope of the DOC, we used a linear mixed-effects model, which explicitly takes into account the fact that those data points in the dissimilarity–overlap plane are not completely independent (because for a data set of  $M$  samples, each sample affects  $(M-1)$  data points). To avoid any potential biases due to the detection of change point, we use data points with overlap larger than the median value, that is 50% of all the data points, for all the data sets (from all the body sites). We repeated this step for 200 bootstrap realizations. The distributions of the slopes for different body sites are shown in Extended Data Fig. 6b. The one-tailed  $P$  values are calculated as the fraction of bootstrap realizations with a non-negative slope, and adjusted for multiple comparisons by the procedure of Benjamini and Hochberg<sup>37</sup>.

We emphasize that those two measures are complementary. In the first measure ( $f_{\text{ns}}$ ), we consider the existence of a negative slope and ask what is the fraction of data points that support it. In the second measure, we consider a fixed fraction of data points (50%) and asked whether a significant negative slope is observed.

**Population dynamics model.** The GLV model represents the dynamics of  $N$  interacting species as a set of ordinary differential equations:

$$\frac{dx_i}{dt} = r_i x_i + \sum_{j=1}^N a_{ij} x_j x_i, \quad i = 1, \dots, N. \quad \text{Here, } r_i \text{ is the intrinsic growth rate of species}$$

$i$ ,  $a_{ij}$  is the interaction strength between species  $j$  and  $i$ , and  $a_{ii}x_i^2$  (with  $a_{ii} < 0$ ) represents the logistic growth term. We considered a microbial 'sample' as a steady state of a GLV model parameterized by the growth rate vector  $\mathbf{r} = \{r_i\} \in \mathbb{R}^N$  and the interaction matrix  $A = (a_{ij}) \in \mathbb{R}^{N \times N}$  and we set  $N = 100$  and  $a_{ii} = -1$  in our simulations. We generated different 'cohorts', each consists of  $M = 100$  'samples'. The GLV models differ from each other in their specific parameters ( $r_i$  and  $a_{ij}$ ). To achieve that, for each cohort, we first constructed a 'base' GLV model ( $\mathbf{r}^*$ ,  $A^*$ ) as follows:  $r_i^*$  is randomly chosen from the uniform distribution  $\mathcal{U}(0, 1)$ .  $a_{ij}^*$  is randomly chosen from the normal distribution  $\mathcal{N}(0, (\sigma_{\max})^2)$ ,  $\sigma_{\max}$  varies between 0 and 1,  $\sigma_{\max}$  is the maximal interaction strength allowed to ensure stability of the ecological system (here  $\sigma_{\max} = 0.1$ ). Then, different GLV models ( $\nu = 1, \dots, M$ ) are generated as random variations of this base model with  $r_i^\nu = \varphi_i^\nu r_i^*$  and  $a_{ij}^\nu = \phi_{ij}^\nu a_{ij}^*$  where both  $\varphi_i^\nu$  and  $\phi_{ij}^\nu$  are randomly chosen from a uniform distribution  $\mathcal{U}(1 - \delta, 1 + \delta)$  so that the expected values of the parameters of those models in the same cohort are the same as the base model, that is,  $E[\mathbf{r}]_\nu = \mathbf{r}^*$  and  $E[A]_\nu = A^*$ . In other words, all the samples of the same cohort are generated from GLV models that share the same structure and sign pattern of the base interaction matrix  $A^*$ . As  $\delta \rightarrow 0$  the model parameters become identical in all the GLV models of the same cohort. Thus,  $\delta \equiv 1 - \delta$  quantifies the 'universality' of the dynamics of those models. Finally, for each cohort, the 100 samples (steady states) were generated by integrating the GLV differential equations with random initial conditions (both initial assemblage and abundance profile are randomly chosen).

**Human microbiome data sets analysed in this work.** *Longitudinal microbiome data sets.* (1) Two time series of gut microbiome consist of 336 and 131 stool samples, respectively. A 16S rRNA gene-based data set, variable region V4, analysed here at the OTU level. For detailed description of this data set see ref. 28. The data are available at <http://qiita.ucsd.edu> under study ID 550. (2) Two time series of gut microbiome consist of 299 and 180 stool samples, respectively. A 16S rRNA gene-based data set, variable region V4, analysed here at the OTU level. For detailed description of this data set see ref. 11. The data are available in the European Bioinformatics Institute (EBI) European Nucleotide Archive (ENA) under the nucleotide accession number ERP006059.

*Cross-sectional microbiome data sets.* To compare quantitatively the universality of microbial dynamics in different body sites, we used two large-scale microbiome data sets. (1) Human Microbiome Project (HMP)<sup>9,15</sup>. A 16S rRNA gene-based data set, variable regions V3 to V5, of the human microbiome from 239 healthy subjects. The data are available at <http://hmpdacc.org/> and are detailed in refs 9, 15. This data set covers 18 body sites in five areas: the oral cavity (nine sites: saliva ( $M = 262$ ), tongue dorsum ( $M = 291$ ), palatine tonsils ( $M = 285$ ), keratinized gingiva ( $M = 289$ ), hard palate ( $M = 275$ ), buccal mucosa ( $M = 287$ ), throat ( $M = 283$ ), and sub- and supragingival plaques ( $M = 283$  and  $M = 289$ , respectively)), the gut (one site: stool ( $M = 297$ )), the vagina (three sites: introitus ( $M = 115$ ), mid-vagina ( $M = 124$ ), and posterior fornix ( $M = 124$ )), the nasal cavity (one site: anterior nares ( $M = 230$ )), and the skin (four sites: left and right antecubital fossae ( $M = 161$  and  $M = 171$ , respectively) and retroauricular creases ( $M = 240$  and  $M = 257$ , respectively)). Full protocols are available on the HMP DACC website (<http://hmpdacc.org/HMMCp>). We performed the DOC analysis at the OUT level. We used a single sample from each subject. In case more than one sample is available, we used the first visit. (2) Student Microbiome Project (SMP). A 16S rRNA gene-based data set, variable region V4 from 85 college-aged adults. The data set covers four body sites: gut ( $M = 72$ ), tongue ( $M = 79$ ), forehead skin ( $M = 78$ ) and palm skin ( $M = 60$ ). In case there are multiple samples measured for one subject, we used the sample

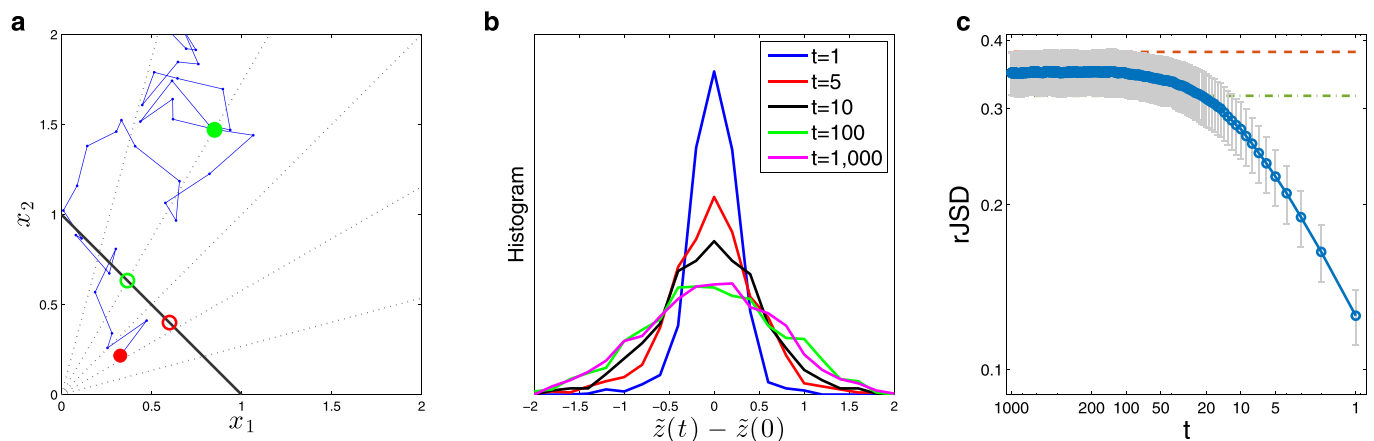
from the first visit. For detailed description of this data set see ref. 16. The data are available at <https://github.com/gregcaporaso/student-microbiome-project/tree/master/otu-tables>.

To rule out several leading candidates of confounding factors in our DOC analysis, we analysed two additional data sets. (3) A data set of healthy volunteers ( $M = 98$ ) from the Cross-sectional Study of Diet and Stool Microbiome Composition (COMBO). Diet information was collected using two questionnaires that queried recent diet (recall) and habitual long-term diet (food frequency questionnaire, FFQ). Stool samples were collected, and DNA samples were analysed by 454/Roche pyrosequencing of the variable region V1–V2 of the 16S rDNA gene segments. We performed the DOC analysis at the OTU level. For detailed description of this data set see ref. 38. (4) A data set of healthy women ( $M = 53$ ), aged 20–55 years (median 42.5), as part of the Flemish Gut Flora Project (FGFP). Stool consistency levels using Bristol stool scale (BSS) scores were self-reported. The V4 region of the 16S rDNA gene was sequenced. We performed the DOC analysis at the OTU level. For detailed description of this data set see ref. 33.

*Clinical trial data set.* Stool samples of patients with rCDI: before and after FMT. This clinical trial was approved by the Partners Human Research Committee as well as by the US Food and Drug Administration (FDA) (Investigational New Drug application number 15199) and registered at ClinicalTrials.gov (NCT01704937). Informed consent was obtained from all participants. Microbial samples from 17 patients with rCDI were analysed in the groups of pre-FMT and post-FMT. Only subjects for whom both pre- and post-FMT samples are available were included. In cases where more than one post-FMT sample is available we included only the first one (median, 4 days after FMT). The V4 region of the 16S rRNA gene was sequenced using an Illumina MiSeq. We performed the DOC analysis at the OTU level. For detailed description of this data set see ref. 17.

*Code availability.* The Matlab code for computing the DOC and the universality measures as well as an example data set (that is, the data set used to generate Fig. 2b) are freely available at the project webpage: <http://scholar.harvard.edu/yyl/doc> and have been added to the Supplementary Information.

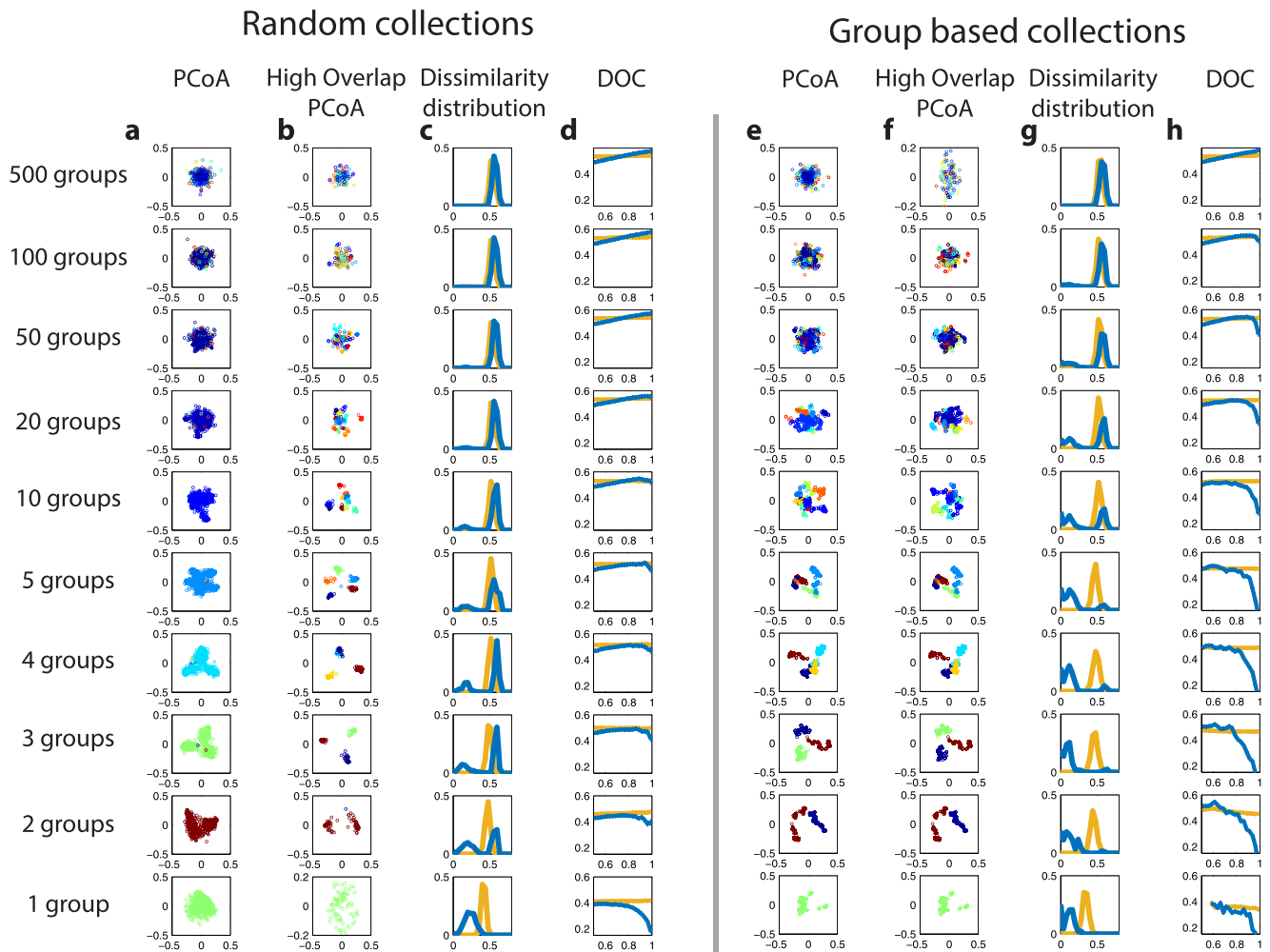
30. Lozupone, C. A., Hamady, M., Kelley, S. T. & Knight, R. Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. *Appl. Environ. Microbiol.* **73**, 1576–1585 (2007).
31. Faith, J. J. *et al.* The long-term stability of the human gut microbiota. *Science* **341**, 1237439 (2013).
32. Gilbert, J. A. & Alverdy, J. Stool consistency as a major confounding factor affecting microbiota composition: an ignored variable? *Gut* **65**, 1–2 (2016).
33. Vandeputte, D. *et al.* Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. *Gut* **65**, 57–62 (2016).
34. Lawley, T. D. *et al.* Targeted restoration of the intestinal microbiota with a simple, defined bacteriotherapy resolves relapsing *Clostridium difficile* disease in mice. *PLoS Pathog.* **8**, e1002995 (2012).
35. Faust, K. & Raes, J. Microbial interactions: from networks to models. *Nat. Rev. Microbiol.* **10**, 538–550 (2012).
36. Goodrich, J. K. *et al.* Conducting a microbiome study. *Cell* **158**, 250–262 (2014).
37. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statist. Soc. B* **57**, 289–300 (1995).
38. Wu, G. D. *et al.* Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**, 105–108 (2011).
39. Bickel, S. L., Tang, K. W. & Grossart, H.-P. Ciliate epibionts associated with crustacean zooplankton in German lakes: distribution, motility and bacterivory. *Front. Microbiol.* **3**, 243 (2012).



**Extended Data Figure 1 | Displacement of normalized  $N$ -dimensional random walks.** **a**, Trajectory of a two-dimensional random-walk represents the absolute abundance of two species  $x_1, x_2$ . The initial state is marked by a red circle and the first 100 steps are shown. The solid black line is the one-dimensional simplex upon which the locations are projected to obtain the relative abundances  $\tilde{x}_1, \tilde{x}_2$ . The dotted lines starting at the origin represent the projection process: all the points in a dotted line have the same relative abundances and they are all projected to the intersection of the dotted line and the simplex (for example, the solid red and green circles are projected to the red and green open circles, respectively). We define a new coordinate  $\tilde{z}(t) \equiv \tilde{x}_2(t) - \tilde{x}_1(t)$  for the location of normalized relative abundance on the simplex. The displacement of the normalized random walk after  $t$  steps is then  $\tilde{z}(t) - \tilde{z}(0)$ , where  $\tilde{z}(0)$  is the projected location of the initial state (see, as an example, the distance

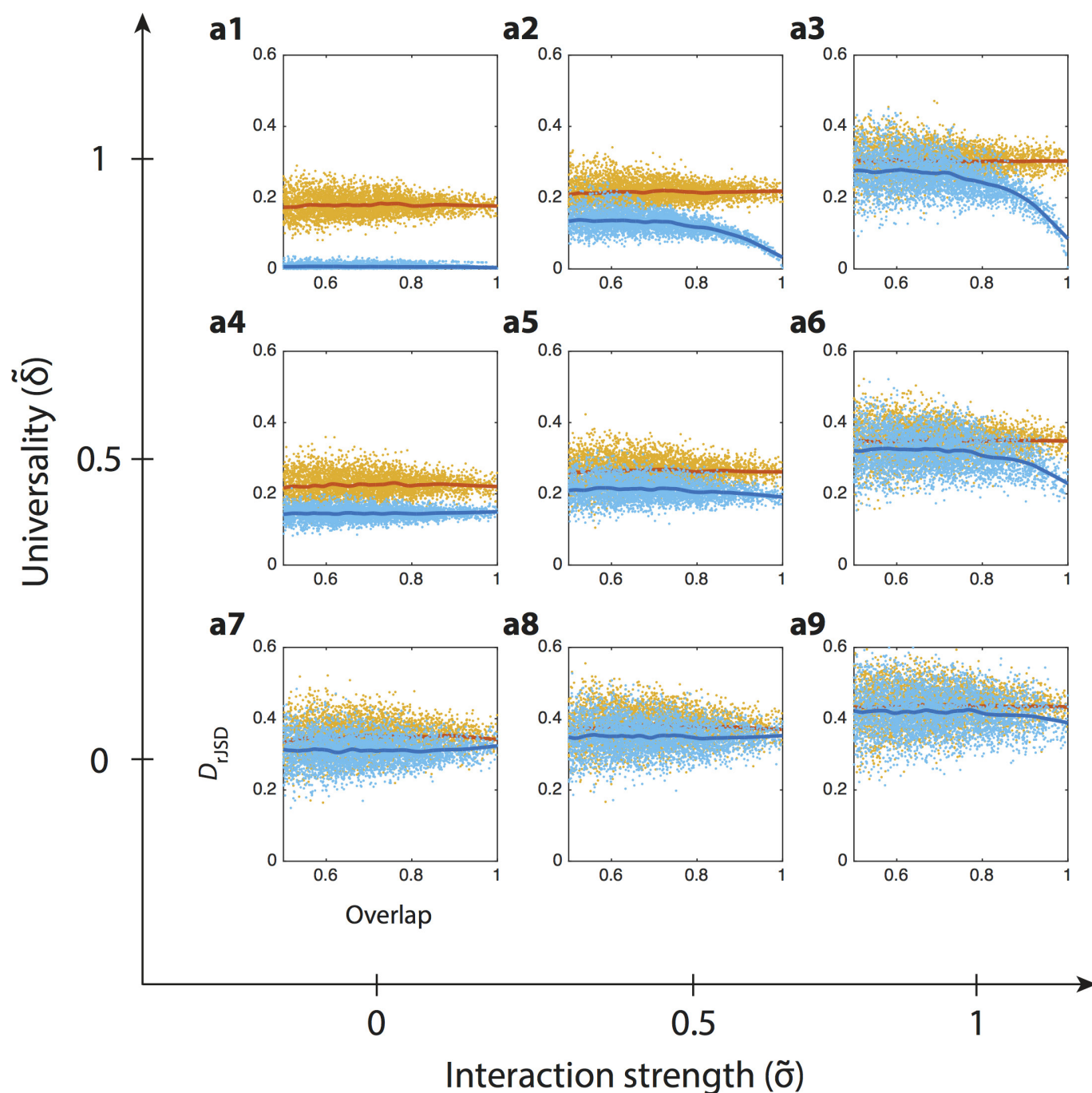
between the green and the red open circles in **a**). **b**, Distributions of displacement of an ensemble of 1,000 random walks after  $t$  steps ( $t = 1, 5, 10, 100, 1,000$ ). For small  $t$ , the displacement distributions depend on  $t$ , while for large  $t$  ( $t = 100, 1,000$ ) the distributions are the same. **c**, Symbols represent the average displacement of 1,000  $N$ -dimensional normalized random walks (here we set  $N = 50$ ), measured as  $D_{\text{rJSD}}$ , and the error bars represent the s.d. Each random walk is forced to stay on the positive orthant, that is, if  $x_i^{(t)} < 0$  we set  $x_i^{(t)} = 0$ .  $D_{\text{rJSD}}$  was calculated using all  $N$  coordinates, setting  $x_i^{(t)} = 10^{-4}$  as a pseudo count for  $x_i^{(t)} = 0$ . Where  $t$  is small, the distance grows with increasing  $t$ ; however, the distance saturates for large  $t$ . The dashed red and green lines represent the average distance between two random locations (green) and between the final locations ( $x^{(t=1,000)}$ ) of the random walks (red).





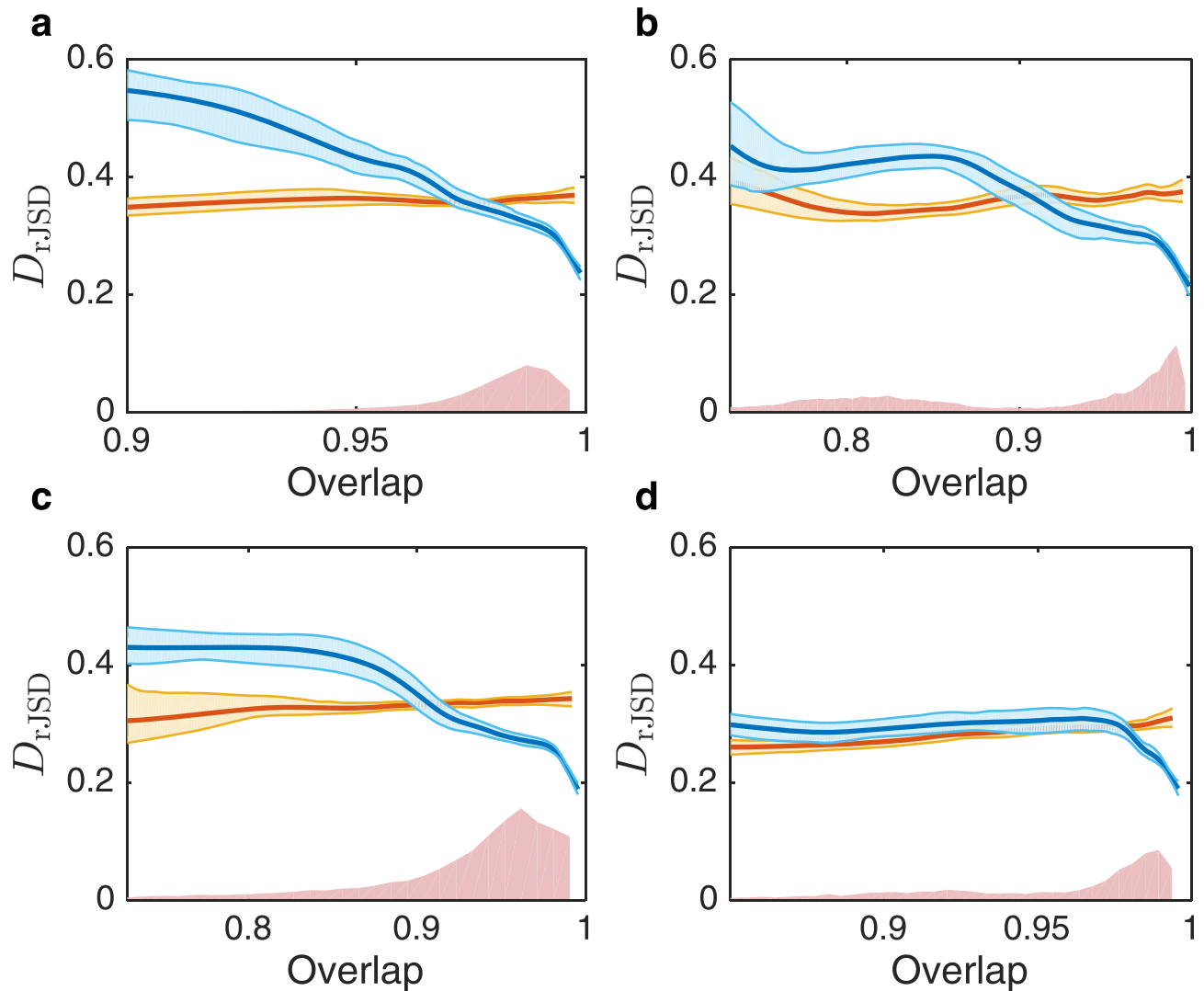
**Extended Data Figure 2 | Detection of group dynamics using an ordination technique.** **a–h**, In each row, 500 synthetic samples were generated. Samples in the same group were taken from the steady states of the same GLV model of 100 species. The initial species assemblages were determined in two scenarios: at random (**a–d**) or on the basis of the group (**e–h**). In the latter scenario, in each group the species were first randomly ordered and then in each of the samples the first  $f$  species were selected and the other been removed ( $f$  is randomly chosen from a uniform distribution  $\mathbb{U}(20, 100)$ ). In columns **a** and **e**, a standard ordination technique, that is, principal coordinate analysis (PCoA), was applied. All 500 samples were shown in the plane of the first two principal coordinates (using rJSD as the distance metric) and coloured according to their group. In **b** and **f**, only the samples that have high overlap ( $>0.95$ ) with at least one other sample were shown. Panels **c** and **g** show the

dissimilarity distributions  $P(rJSD)$  between the high-overlap sample pairs. Panels **d** and **h** show the DOCs. The ordination technique successfully detects the existence of group dynamics (especially when the number of groups is small). We anticipate that the group dynamics can also be detected by classical clustering analysis. In the scenario of random collections, the PCoA of high-overlap samples, that is, samples that have high overlap ( $>0.95$ ) with at least one other sample, is doing better than the PCoA of all samples to detect group dynamics, especially for a small number ( $\sim 2$ – $10$ ) of groups. Moreover, for a small number of groups, the dissimilarity distributions  $P(rJSD)$  can distinguish between the two scenarios of initial assemblage selection: random or group-based. The ordination technique cannot distinguish between the cases of 500 groups (individual dynamics) and single group (universal dynamics). Those cases can be distinguished by the DOC analysis.



**Extended Data Figure 3 | Detecting universality in population dynamics models.** Synthetic microbial samples were calculated as steady states of GLV models (see Methods). The GLV models are generated as cohorts (100 models in each cohort) with different levels of (i) inter-species interaction strength; and (ii) universality, tuned by the parameters  $\tilde{\sigma}$  and  $\tilde{\delta}$ , respectively (see Methods). In each of the 100 models, a random fraction  $f$  of the species ( $f \sim \mathbb{U}(0, 0.8)$ ) was initially removed, and the remaining species were initiated with random abundance ( $x \sim \mathbb{U}(0, 1)$ ). The dissimilarity–overlap points of sample pairs in each cohort and of the corresponding randomized samples are shown in light blue and yellow, respectively. The solid curves represent the DOCs calculated using the

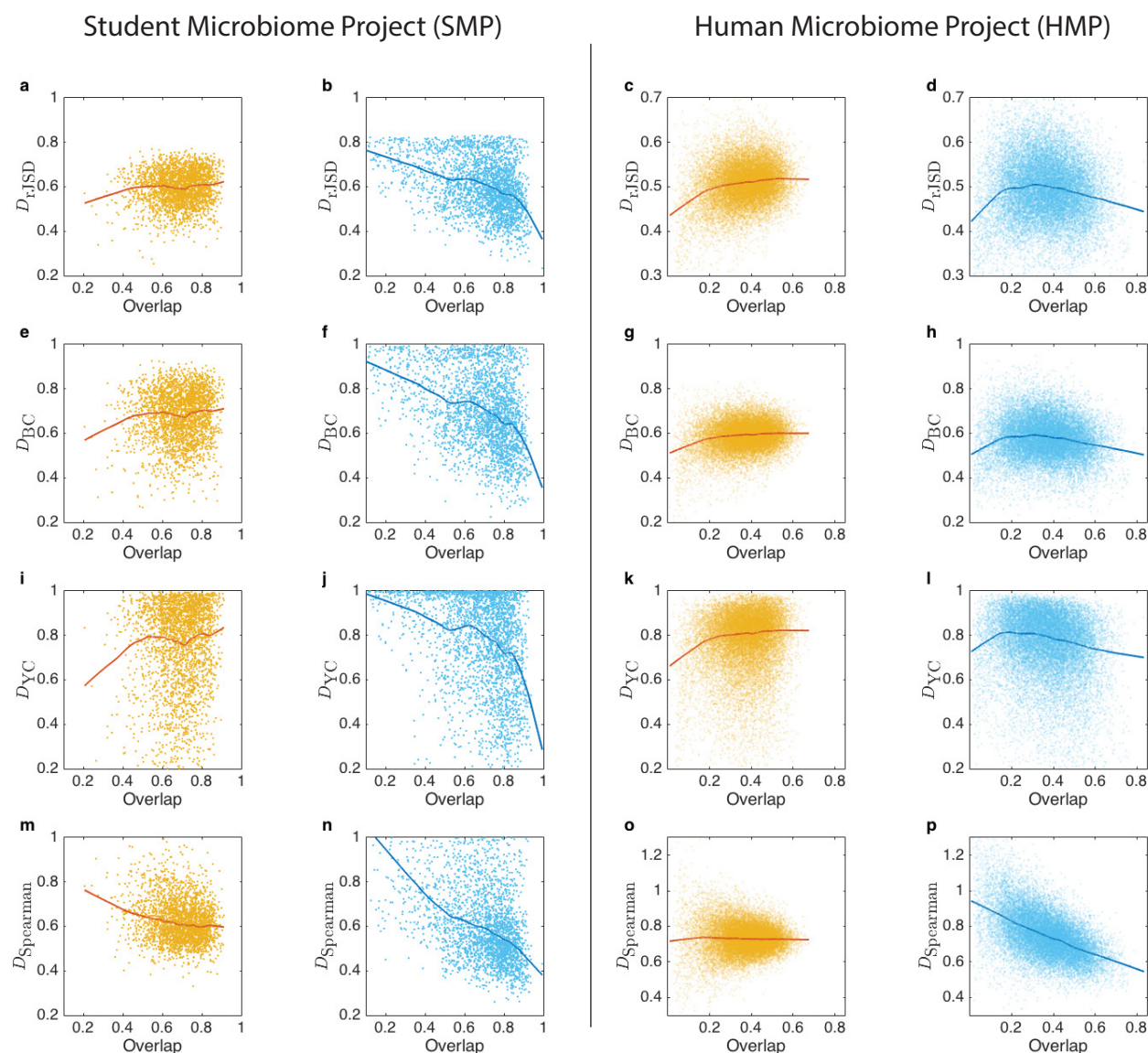
robust LOWESS method. The DOC of cohorts generated by GLV models without inter-species interactions (**a1**, **a4**, **a7**) is flat even in the high-overlap region. This is because, without inter-species interactions, for any sample pair the presence or absence of unique (that is, non-shared) species has no effect on the shared ones. A flat DOC is also observed in the case of individual dynamics (**a7**, **a8**, **a9**), where a higher overlap between sample pairs does not lead to more similar abundance profiles. However, in the case of universal dynamics with strong inter-species interactions (for example, **a3**), the DOC displays a clear negative slope in the high-overlap region.



**Extended Data Figure 4 | DOC analysis of gut microbiome samples from longitudinal studies.** **a–d**, Sample pairs are selected from four different subjects, with number of samples:  $M_a = 299$ ,  $M_b = 180$ ,  $M_c = 336$ ,  $M_d = 131$ , respectively. The mean DOCs (calculated from 100 bootstrap realizations using the robust LOWESS method) of each subject and the corresponding randomized samples are shown in dark blue and yellow, respectively. The shaded area indicates the range of the 94% confidence intervals. The overlap distributions are shown in red. For all the four subjects, a clear negative slope of the DOC is observed at the high-overlap

region, indicating largely time-invariant or universal dynamics for each subject throughout the measurement period. This is in marked contrast with the flat DOC of the null model (see Supplementary Information section 1.3). The secondary peak of lower-overlap samples in **b** (overlap of ~0.8) is of sample pairs from two different periods, before and after a *Salmonella* infection, which represent two distinct microbial steady states and thus exhibit a flat DOC. This is consistent with our assumption of time-invariant microbial dynamics for a given healthy individual.

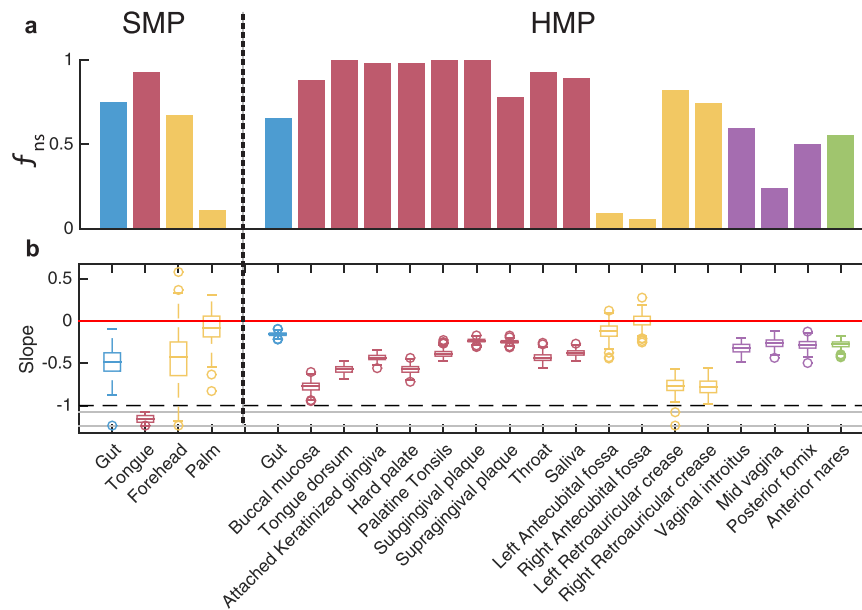




**Extended Data Figure 5 | DOC analysis of gut microbiome samples is consistent across different studies and different dissimilarity measures.**

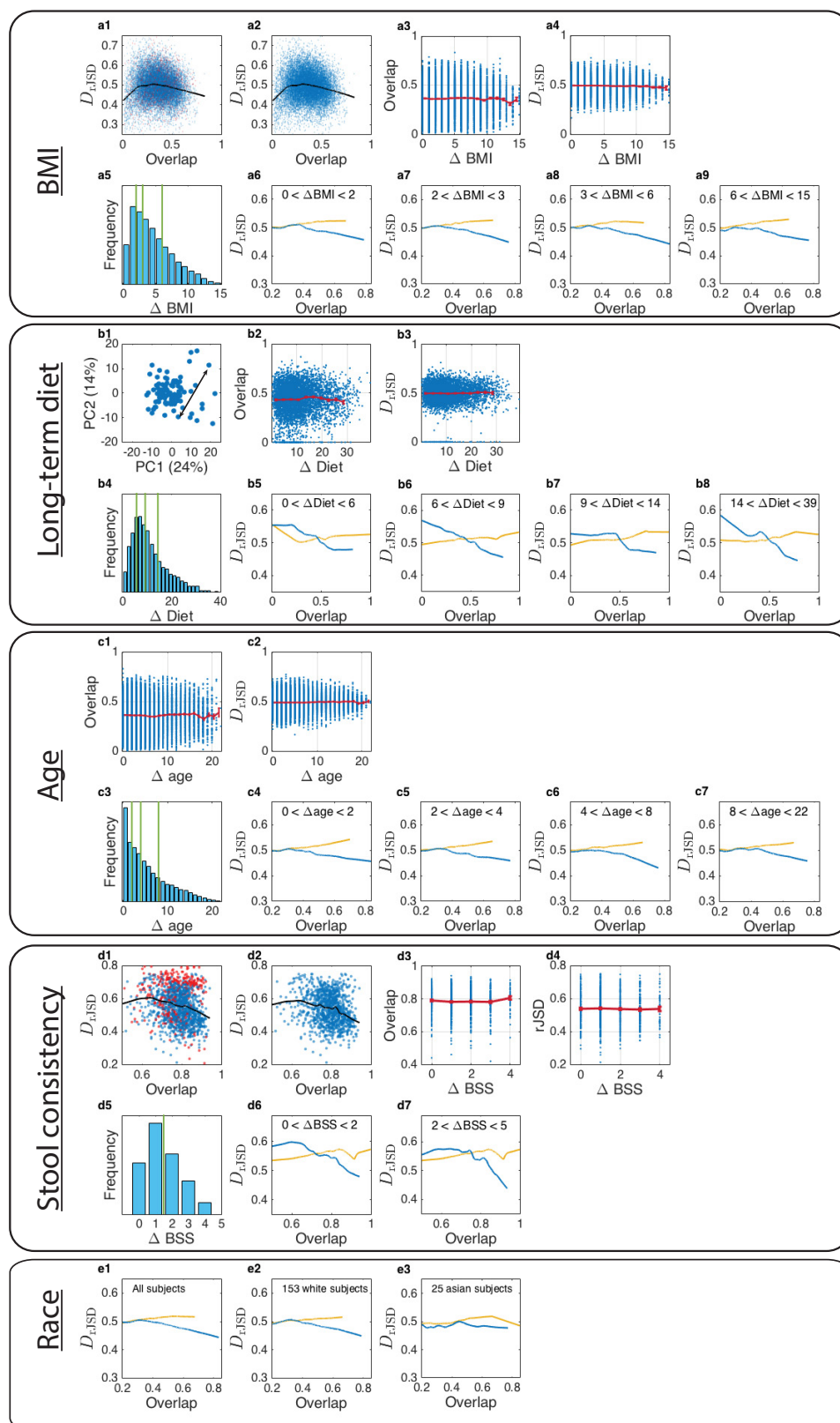
For two microbiome samples, the dissimilarity of their abundance profiles over shared species can be evaluated by different measures. Weighted measures, such as rJSD, Bray–Curtis (BC) dissimilarity and Yue–Clayton (YC) dissimilarity should be applied to the renormalized abundance profiles, to ensure mathematical independence between the overlap and the dissimilarity measures. Rank-based dissimilarity measures, for

example, negative Spearman correlation (nSC), can be directly applied without renormalization. We used the four dissimilarity measures (rJSD, BC, YC and nSC) to calculate the DOC (using robust LOWESS) of gut microbiome samples from two studies: HMP and SMP. In all cases, we observed a pronounced negative slope in the DOC (dark-blue curve) of real sample pairs (light-blue points) and a flat DOC (orange curve) for the pairs of randomized samples (yellow points).



**Extended Data Figure 6 | Quantifying the universality of human microbial dynamics in different body sites.** **a**, The fraction ( $f_{ns}$ ) of data for which a negative slope is observed in Fig. 3. Note that for overlap values close to zero (for example, Fig. 3d, f1–4, g1–3) a positive slope occurs as the artefact of dissimilarity between relative abundance profiles with small number of species (see Supplementary Information section 1.1.3). For gut and mouth, a negative slope of DOC is observed in the two data sets for a broad range of overlap, indicating a significant universality of microbial dynamics in those habitats. By contrast, the negative slope of DOC in the hand's skin microbiome is observed only for a small part of the sample

pairs. **b**, Box plot of the slope of DOC calculated from 200 bootstrap realizations. The slope is calculated by fitting a linear mixed-effects model for data points with overlap larger than the median. We report one-tailed  $P$  values, calculated as the fraction of bootstrap realizations with a non-negative slope, adjusted for multiple comparisons by the procedure of Benjamini and Hochberg. The null hypothesis of non-negative slope is rejected for all body sites ( $P < 1 \times 10^{-2}$ ) except four skin sites: forehead ( $P = 0.099$ ), palm ( $P = 0.377$ ) in the SMP study and left/right antecubital fossa in the HMP study ( $P = 0.099$  and  $P = 0.495$ ).



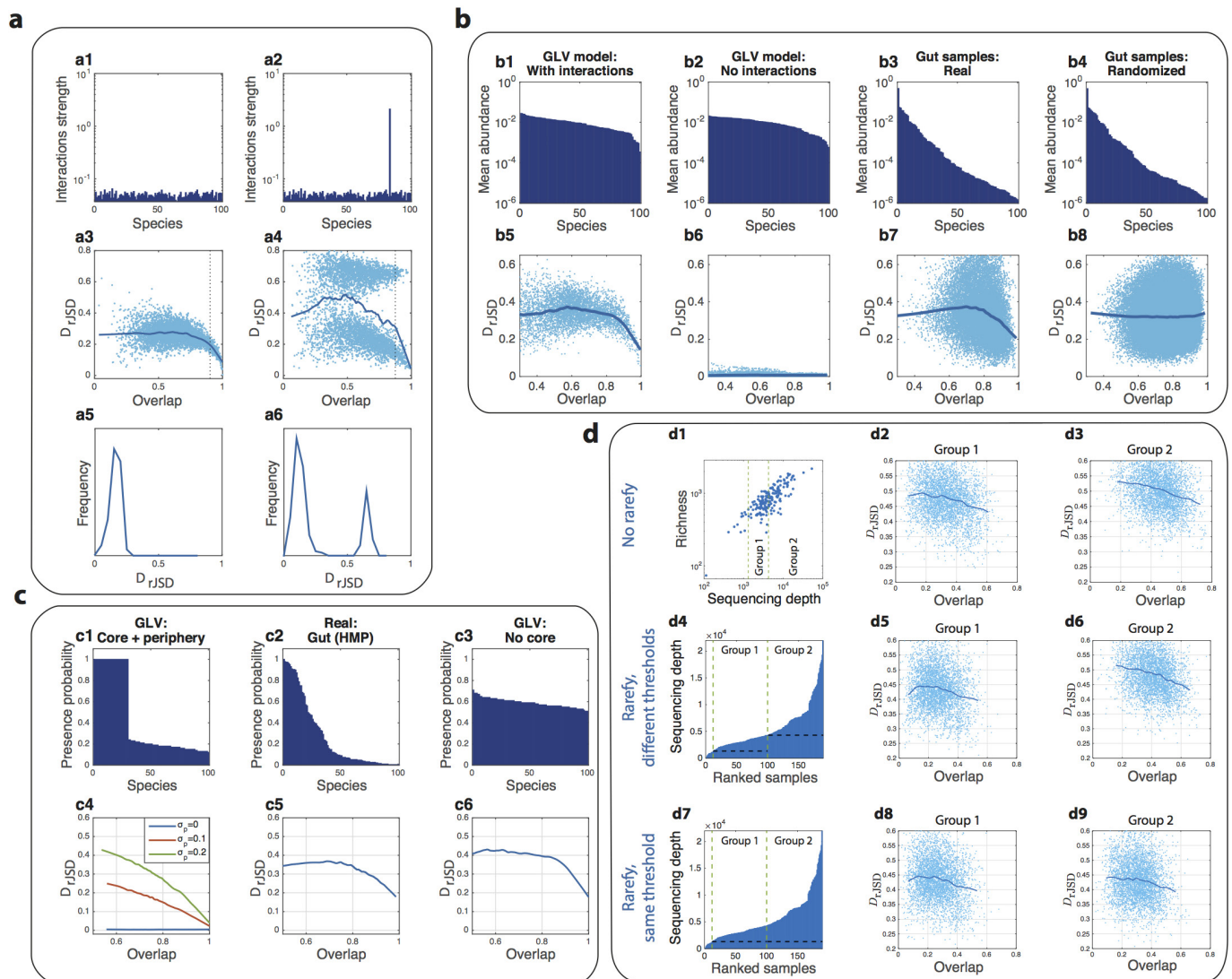
Extended Data Figure 7 | See next page for caption.



# Extended Data Figure 7 | Effects of various host factors on the DOC analysis.

**a**, The effect of body mass index (BMI) on the DOC analysis. **a1**, DOC analysis of all gut microbiome sample pairs among 190 subjects from the HMP study. Red points represent samples pairs associated with at least one obese subject (with BMI > 30). **a2**, Same as in **a1**, but 13 obese subjects with BMI > 30 were excluded. **a3**, Blue points represent the gut microbiome samples' overlap and  $\Delta$ BMI. The red curve is the average (error bars represent the s.e.m.). **a4**, Dissimilarity versus  $\Delta$ BMI. **a5**, Distribution of  $\Delta$ BMI values, divided into four groups of equal number of pairs. **a6–a9**, DOC analysis of the sample pairs in each group. **b**, The effect of diet on the DOC analysis. **b1**, Diet difference ( $\Delta$ diet) between two subjects is defined as the Euclidean distance between their associated diet scores in the two leading principal components PC1 and PC2. In total there are  $M=97$  healthy subjects in the COMBO study<sup>38</sup>. **b2**, Overlap versus  $\Delta$ diet. Blue points represent the overlap and  $\Delta$ diet of all gut microbiome pairs among the 97 subjects from the COMBO study. The red curve is the average (error bars represent the s.e.m.). **b3**, Dissimilarity versus  $\Delta$ diet. **b4**, Distribution of  $\Delta$ diet values, divided into four groups of equal number of pairs. **b5–b8**, DOC analysis of the pairs in each group. **c**, The effect of age on the DOC analysis. **c1**, Overlap

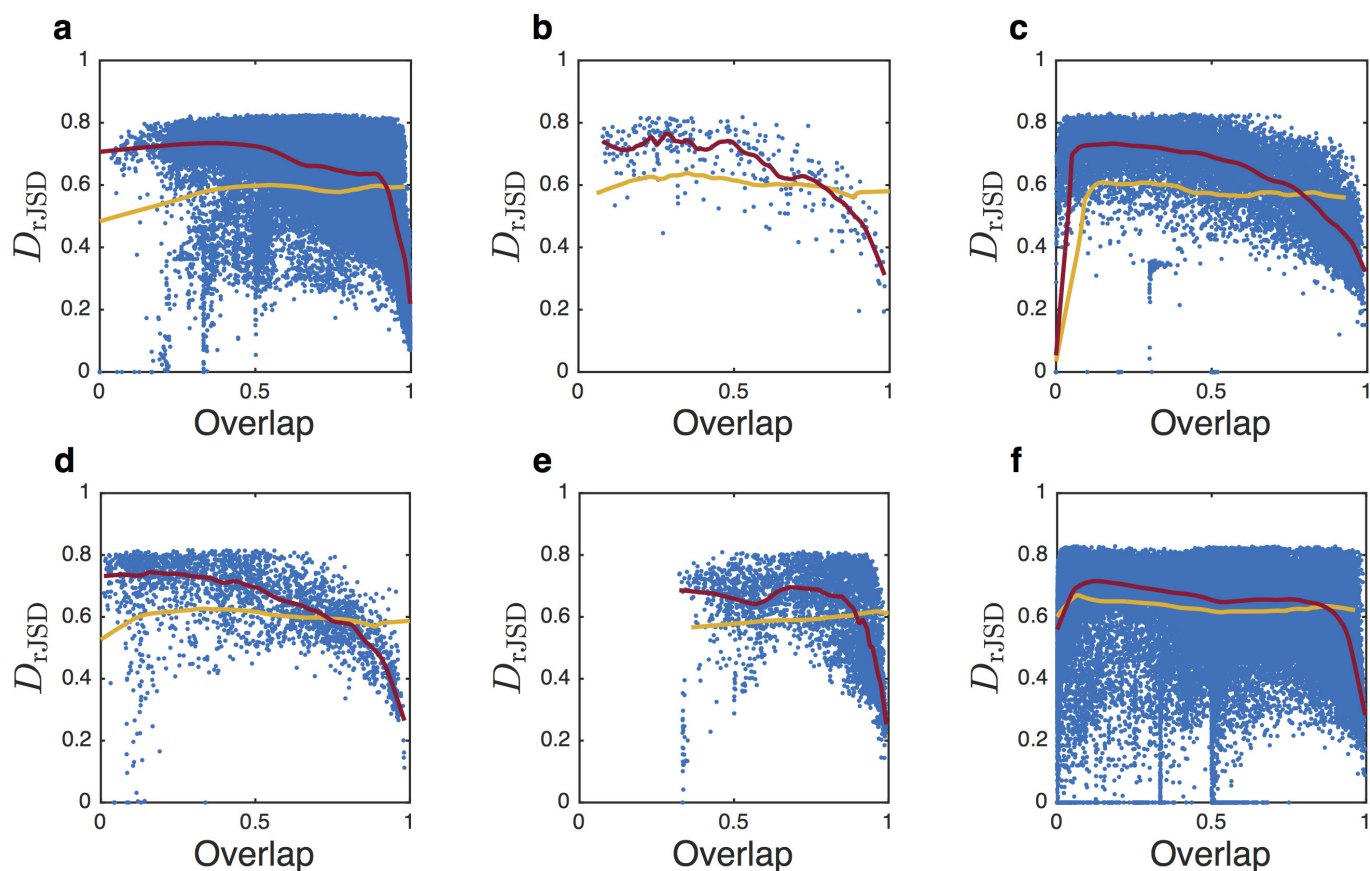
versus  $\Delta$ diet. Blue points represent the overlap and  $\Delta$ age of all gut microbiome samples pairs between the 190 subjects from the HMP study. The red curve is the average (error bars represent the s.e.m.). **c2**, Dissimilarity versus  $\Delta$ age. **c3**, Distribution of  $\Delta$ age values, divided into four groups of equal number of pairs. **c4–c7**, DOC analysis of the pairs in each group. **d**, The effect of stool consistency on the DOC analysis. **d1**, DOC analysis of all sample pairs. In this data set the subjects have BSS values between 1 and 6. The points (sample pairs) associated with subjects with BSS = 6 (at least one subject has BSS = 6) are coloured in red. The black line is the DOC. **d2**, DOC analysis of all subjects with BSS < 6. **d3, d4**, Among all subjects with  $1 \leq \text{BSS} \leq 5$ , the overlap and the dissimilarity are independent of  $\Delta$ BSS. **d5**, Distribution of  $\Delta$ BSS values for the 46 subjects with  $1 \leq \text{BSS} \leq 5$ . **d6, d7**, DOC analysis of the pairs with similar BSS values,  $0 \leq \Delta \text{BSS} \leq 1$  (**d6**), and pairs with more different BSS values,  $2 \leq \Delta \text{BSS} \leq 4$  (**d7**). In both cases, a clear negative slope of the DOC is observed. **e**, The effect of race on the DOC analysis. **e1**, All subjects ( $M=190$ ). **e2, e3**, White subjects ( $M=153$ ) (**e2**) and Asian subjects ( $M=25$ ) (**e3**). Note that in the HMP study, stool samples were collected from 153 white subjects, 10 black subjects, 25 Asian subjects, and 2 subjects from other races.



### Extended Data Figure 8 | DOC analysis under special conditions.

**a**, The effect of strongly interacting species. A comparison of two GLV models of 100 species with random inter-species interactions. The system parameters were fixed for all the simulated samples ( $M = 100$ ), representing maximal universality. In **a1**, all species have the same characteristic interaction strength, while in **a2**, the inter-species interactions of one species are markedly stronger than all other species, representing a strongly interacting species. The presence/absence of the strongly interacting species markedly affects (either directly or indirectly) the abundance profile of many other species, leading to a pronounced secondary cloud of points in the dissimilarity–overlap plane (**a4**). The effect is the most pronounced in the region of high-overlap (top 5%) pairs, and can be detected by looking at their dissimilarity distributions (**a5**, **a6**). **b**, DOC behaves the same for samples with uniform or skewed abundance distribution. **b1**, **b2**, Samples were generated from the steady states of the GLV model with largely uniform abundance distribution (determined mainly by the species growth rates). In the presence of inter-species interactions (**b1**), a negative slope of the DOC is observed. By contrast, in the absence of inter-species interactions (**b2**), a flat DOC is observed. **b3**, Real samples from the gut (from the HMP study, genus level) exhibit a high level of alpha-diversity and a very skewed abundance distribution. A negative slope of the DOC in the high-overlap region is observed. **b4**, The randomized samples preserve the abundance distribution of the real samples but the effect of inter-species interactions is removed, leading to a flat DOC. **c**, Effect of core species and non-interacting periphery species. **c1**, Samples were generated as steady states of the GLV model with  $N = 100$  species. The parameters of the GLV model were fixed for all the samples, representing maximal universality. The initial

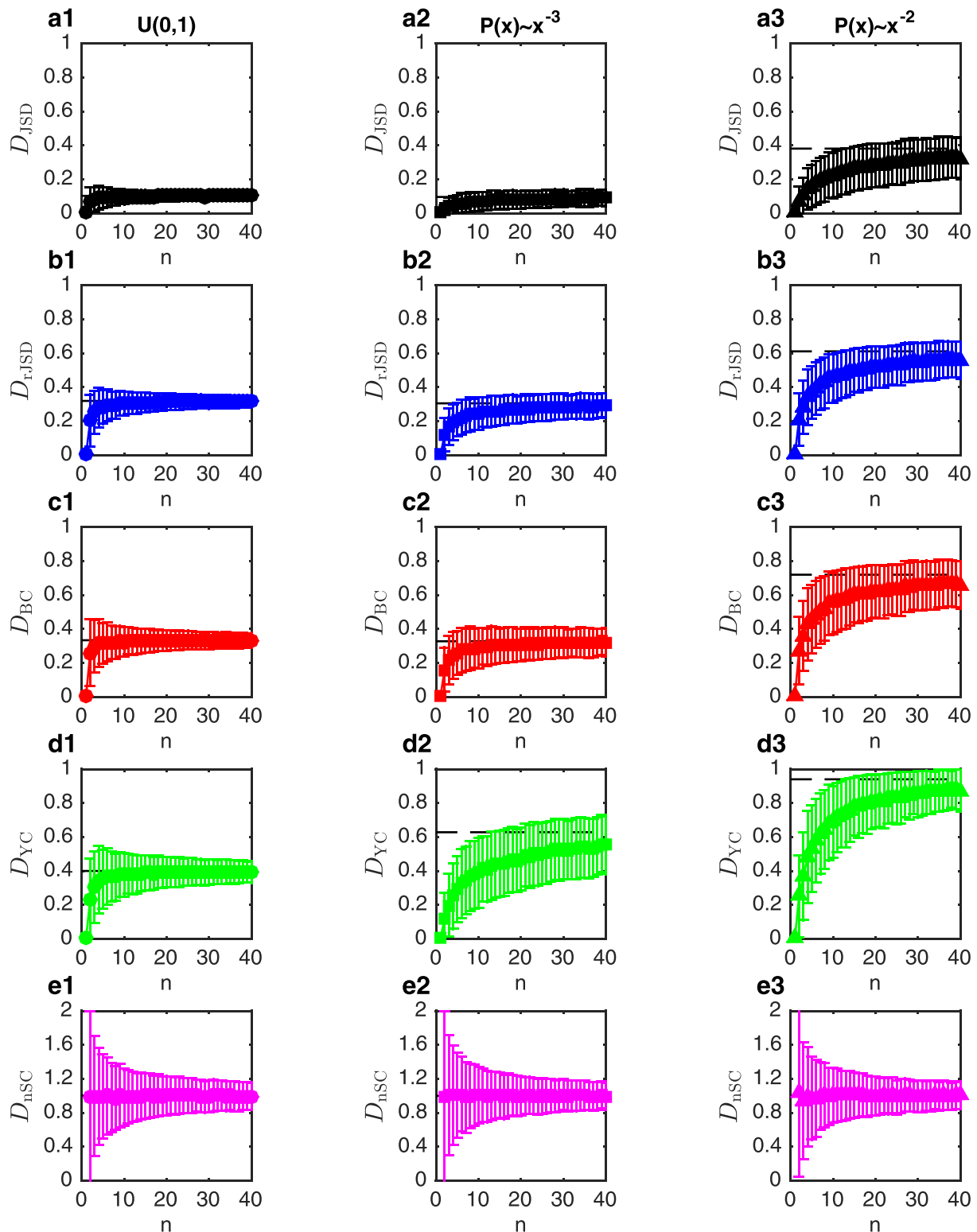
species assemblages were chosen as follows: 30 species were present in all the samples, representing a set of ‘core species’, and the other 70 ‘periphery’ species were present with lower probability (mean 0.18, min 0.12, and max 0.24). **c2**, Presence probability of real gut microbial samples, from the HMP at the genus level. Only one genus (*Bacterioides*) is present in all the samples. **c3**, Species presence probability in a GLV model where all species are present with average probability 0.6. **c4**, The effect of the interactions of the peripheral species. In the GLV model, the inter-species interactions among the core species (core–core) has a characteristic strength  $\sigma_{\text{core}} = 0.15$ , and both the periphery–periphery and the periphery–core interactions have a characteristic strength  $\sigma_p$ . When  $\sigma_p = 0$ , that is, the peripheral species do not interact with the core species, the DOC is flat. When  $\sigma_p > 0$ , the DOC has a negative slope. **c5**, **c6**, In the case of real gut microbiome samples as well as the GLV model without core species, the DOC has a negative slope in the high-overlap region. **d**, The effect of sequencing depth on the DOC analysis. **d1**, Richness (number of present OTUs) versus sequencing depth of 190 HMP gut samples. 12 subjects with fewer than 1,300 reads per sample were excluded and the remaining 178 were assigned into two groups of  $n = 89$  subjects, with average sequencing depth 3,019 and 8,640 reads per sample. **d2**, **d3**, The characteristic overlap between samples of group 1 is smaller than between samples of group 2. However, DOC analysis of each group shows a clear negative slope. **d4–d6**, Samples of each group were rarefied before analysis with minimal community size of 1,317 and 4,333 in group 1 and 2, respectively, as represented by the black dashed lines in **d4**. **d7–d9**, Samples of both groups were rarefied before analysis with the same minimal community size of 1,317, as represented by the black dashed line in **d7**.



**Extended Data Figure 9 | DOC analysis of longitudinal microbiome data from six lakes in Germany<sup>39</sup>.** Data downloaded from <http://qiita.microbio.me>, study ID 945. **a**, Stechlin ( $M = 440$ ). **b**, Haus ( $M = 26$ ). **c**, Tiefwaren ( $M = 164$ ). **d**, Melzer ( $M = 68$ ). **e**, Breiter Luzin ( $M = 89$ ). **f**, Fuchskuhle ( $M = 355$ ). Blue points represent the dissimilarity–overlap values of sample pairs from the same lake. The DOCs of real samples from each lake and that from the corresponding randomized samples are calculated using robust LOWESS and shown in red and yellow,

respectively. For all the six lakes, a clear negative slope is observed for the DOCs of real samples, suggesting universal or time-invariant microbial dynamics for each lake. Differences in the DOC shapes (for example, the moderate DOC slope in **b**, **c** and **d**, in contrast with the steep DOC in **a**, **e** and **f**) deserve a systematic study of those microbial ecosystems. This example clearly demonstrates the applicability of DOC analysis to general microbial ecosystems, for example, soil, ocean, rhizosphere/phylosphere and fermenters.





**Extended Data Figure 10 | Average dissimilarity between two normalized random vectors.** Two independent vectors  $\mathbf{x}, \mathbf{y}$  of  $n$  elements randomly chosen from the uniform distribution  $\mathbb{U}(0, 1)$  were generated and then normalized  $\hat{\mathbf{x}}_i \equiv \frac{x_i}{\sum_{j=1}^n x_j}$  and  $\hat{\mathbf{y}}_i \equiv \frac{y_i}{\sum_{j=1}^n y_j}$ . (Note that in practice all  $n$  elements are always shared in  $\mathbf{x}$  and  $\mathbf{y}$ , since zeros are very unlikely.) The dissimilarity  $D(\hat{\mathbf{x}}, \hat{\mathbf{y}})$  is then calculated using the five dissimilarity measures ( $D_{\text{JSD}}, D_{\text{rJSD}}, D_{\text{BC}}, D_{\text{YC}}$  and  $D_{\text{nSC}}$ ). Average dissimilarity and standard deviations of 1,000 pairs are shown in **a1, b1, c1, d1** and **e1**,

for the different measures. The horizontal black dashed line represents the average dissimilarity for  $n = 100$ . For all the measures here, the dissimilarity displays no  $n$ -dependence for  $n > 15$ , while  $D_{\text{nSC}}$  is  $n$ -independent for any  $n > 0$ . Similar analysis was performed for vectors whose elements were chosen from power-law distributions  $P(x) \sim x^{-\alpha}$  with  $\alpha = 3$  (**a2, b2, c2, d2** and **e2**) and  $P(x) \sim x^{-\alpha}$  with  $\alpha = 2$  (**a3, b3, c3, d3** and **e3**).

# Development of the gut microbiota and mucosal IgA responses in twins and gnotobiotic mice

Joseph D. Planer<sup>1,2</sup>, Yangqing Peng<sup>1,2</sup>, Andrew L. Kau<sup>1,2</sup>, Laura V. Blanton<sup>1,2</sup>, I. Malick Ndao<sup>3</sup>, Phillip I. Tarr<sup>3</sup>, Barbara B. Warner<sup>3</sup> & Jeffrey I. Gordon<sup>1,2</sup>

Immunoglobulin A (IgA), the major class of antibody secreted by the gut mucosa, is an important contributor to gut barrier function<sup>1–3</sup>. The repertoire of IgA bound to gut bacteria reflects both T-cell-dependent and -independent pathways<sup>4,5</sup>, plus glycans present on the antibody's secretory component<sup>6</sup>. Human gut bacterial taxa targeted by IgA in the setting of barrier dysfunction are capable of producing intestinal pathology when isolated and transferred to gnotobiotic mice<sup>7,8</sup>. A complex reorientation of gut immunity occurs as infants transition from passively acquired IgA present in breast milk to host-derived IgA<sup>9–11</sup>. How IgA responses co-develop with assembly of the microbiota during this period remains poorly understood. Here, we (1) identify a set of age-discriminatory bacterial taxa whose representations define a program of microbiota assembly and maturation during the first 2 postnatal years that is shared across 40 healthy twin pairs in the USA; (2) describe a pattern of progression of gut mucosal IgA responses to bacterial members of the microbiota that is highly distinctive for family members (twin pairs) during the first several postnatal months then generalizes across pairs in the second year; and (3) assess the effects of zygosity, birth mode, and breast feeding. Age-associated differences in these IgA responses can be recapitulated in young germ-free mice, colonized with faecal microbiota obtained from two twin pairs at 6 and 18 months of age, and fed a sequence of human diets that simulate the transition from milk feeding to complementary foods. Most of these responses were robust to diet, suggesting that 'intrinsic' properties of community members play a dominant role in dictating IgA responses. The approach described can be used to define gut mucosal immune development in health and disease states and to help discover ways of repairing or preventing perturbations in this facet of host immunity.

To define the relationship between assembly of the gut community and gut mucosal IgA responses, we collected faecal samples monthly for the first 24–36 months of postnatal life from each member of a birth cohort of 40 twin pairs (21 monozygotic) who lived in the greater metropolitan area of a single city in the USA (St. Louis, Missouri). All twins had healthy growth phenotypes as judged by serial anthropometry; 13 pairs were delivered vaginally, 24 by Caesarean section, and three pairs were discordant for mode of birth; 96% received breast milk, infant formula, or a combination of the two as the predominant food source throughout the first 6 months of postnatal life (Supplementary Tables 1–4).

Gut microbiota assembly was defined following an approach based on our previous studies of healthy Bangladeshi and Malawian infants and children<sup>12,13</sup>. We generated a random forests (RF)-derived model of microbiota development from a bacterial V4–16S rRNA data set generated from 1,670 faecal samples collected from the 40 twin pairs ( $20.9 \pm 6.2$  (mean  $\pm$  s.d.) samples per individual). The sparse RF-generated model, based on the relative abundances of the 25 most age-discriminatory operational taxonomic units (OTUs), could predict

chronological age for members of twin pairs as well as for biologically unrelated individuals (OTUs defined by mapping sequenced reads to a reference database of 16S rRNA sequences; see Methods, Extended Data Figs 1 and 2 and Supplementary Tables 5–8). We then conducted a series of reciprocal tests with the data sets we generated from the three birth cohorts. We applied each sparse model to the population of healthy infants and children from which it was generated as well as to the other two populations. We found that the USA model performed consistently across the three populations (Spearman's correlation coefficients of 0.73 and 0.78 for the Bangladeshi and Malawian data sets, respectively; see Methods and Supplementary Table 9).

Although previous studies have identified taxa that are shared more commonly between adult monozygotic compared with dizygotic twin pairs<sup>14,15</sup>, our analysis indicated that none of the 25 age-discriminatory OTUs showed significantly greater concordance in their relative abundances in monozygotic compared with dizygotic twin pairs (Supplementary Table 10). The impact of age, family, milk feeding history, and birth mode on the overall phylogenetic configuration of the microbiota was evaluated with a permutational multivariate analysis of variance (PERMANOVA) and the UniFrac metric. Family had the largest effect (Extended Data Fig. 3), followed by age, and milk feeding (that is, breast milk versus formula) (36%, 11%, and 1%, respectively, when considering only those samples with associated feeding data;  $P < 0.001$  for all variables except birth mode, which did not have a significant effect). A previous study, conducted in the immediate postpartum period, reported that infants born by Caesarean section have a greater representation of skin-derived taxa than those that were vaginally delivered<sup>16</sup>. A caveat to our study is that we were not able to determine the very early effects of birth mode since the median time point for first faecal sampling was postpartum day 52.

Faecal biospecimens were categorized as obtained from donors who were 'predominantly formula fed' or 'predominantly breast fed' at the time of sampling ('predominant' defined as comprising  $\geq 50\%$  of that individual's total milk feedings; Extended Data Fig. 4a and Supplementary Table 2). Linear mixed-effects modelling disclosed that milk feeding practice had a significant effect on maturity ( $P < 0.001$ , ANOVA with predicted microbiota age as the dependent variable and individual/family/chronological age as nested effects). In a post-hoc analysis, infants receiving  $> 50\%$  of their milk from formula feedings had significantly accelerated development of their microbiota during the first 6–7 months of postnatal life compared with infants receiving most of their milk from breastfeeding ( $n = 619$  and 127 samples, respectively; Mann–Whitney  $U$ -test). These differences were no longer statistically significant by 12 months (Extended Data Fig. 4b). This finding can be explained in part by the significantly lower aggregate relative abundance of members of the genus *Bifidobacterium* represented in the RF model in the faecal microbiota of formula-fed infants (Extended Data Fig. 4c, Supplementary Table 11 and ref. 17).

<sup>1</sup>Center for Genome Sciences & Systems Biology, Washington University School of Medicine, St. Louis, Missouri 63110, USA. <sup>2</sup>Center for Gut Microbiome and Nutrition Research, Washington University School of Medicine, St. Louis, Missouri 63110, USA. <sup>3</sup>Department of Pediatrics, Washington University School of Medicine, St. Louis, Missouri 63110, USA.

Faecal samples collected during the first postnatal month, and at 3-month intervals thereafter from each member of the 40 twin pairs, were subjected to fluorescence-activated cell sorting (FACS) to characterize the patterns of IgA targeting of bacterial taxa in their developing microbiota (Supplementary Table 12; see Methods for a description of 'BugFACS' with anti-human IgA). V4-16S rRNA gene sequencing was performed on three fractions generated from each sample ('input', IgA<sup>+</sup>, and IgA<sup>-</sup>). The differential representation of a given taxon between the IgA<sup>+</sup> and IgA<sup>-</sup> fractions is expressed in the form of a log-normalized 'IgA index' that ranges, in theory, from -1 to 1, with positive and negative values indicating enrichment in the IgA<sup>+</sup> and IgA<sup>-</sup> fraction, respectively<sup>8</sup> (Fig. 1a). IgA indices are not a simple reflection of the relative abundances of organisms in the input fraction (Extended Data Fig. 5a).

We identified 30 OTUs that were significantly enriched in either the IgA<sup>+</sup> or IgA<sup>-</sup> fraction in three or more age bins (Fig. 1a). Seven OTUs exhibited consistently positive IgA indices after the third month of life, including two age-discriminatory members of the sparse RF-generated model of gut microbiota development (*Clostridium nexile* OTU 4436046, *Bifidobacterium bifidum* OTU 365385; Fig. 1a). Seventeen OTUs remained untargeted throughout the first 24 months, including six OTUs in the RF-based model (Fig. 1a). Two OTUs manifested significant differences in their IgA targeting during the first 2 postnatal years: *B. longum* (OTU C.1) and *Escherichia coli* (OTU C.3) (Extended Data Fig. 5b and Supplementary Table 13).

We performed an indicator species analysis<sup>18</sup> across all time points to obtain a metric complementary to the IgA index that could describe the strength of partitioning of the 30 OTUs into the IgA<sup>+</sup> or IgA<sup>-</sup> fractions. The results were largely concordant with those obtained from the IgA index-based analysis and provided an additional level of resolution of the temporal patterns and specificity of targeting (Supplementary Table 14 and Extended Data Fig. 6).

IgA indices were highly correlated within twin pairs during the first 21 months of life (Fig. 1b). Indices between unrelated infants were very weakly correlated during the first 6 postnatal months, became increasingly more correlated during the second year of life, and by 24 months co-twins no longer had an IgA response that was significantly more similar to one another than to other unrelated children (Wilcoxon signed-rank test; Fig. 1b). As the effects of family membership diminished, variation of the IgA index for a given taxon across the population of twins also diminished (Extended Data Fig. 5c). The similarity in the IgA profiles between mothers sampled during the first 12 postpartum months (39 mothers;  $3.0 \pm 1.0$  (mean  $\pm$  s.d.) samples per mother) and children at 24 months of life supports the notion that development of a child's gut mucosal IgA responses reaches a state of maturation that resembles that of adults by this age (Supplementary Tables 15 and 16 and Fig. 1c).

On the basis of Pearson's correlation distance, we determined that age and family membership explained the most variance in IgA indices (25% and 19%, respectively), while zygosity and mode of delivery had small but statistically significant effects (0.6% and 0.5%, respectively; PERMANOVA with 999 permutations). Breastfeeding explained 5% of the variance in the model ( $P < 0.001$  for breast milk versus formula feeding as well as for age and family). Intriguingly, IgA targeting of two taxa, *E. coli* (OTU C.3) and *Ruminococcus gnavus* (OTU C.4), varied between children who were predominantly breastfed and those who were predominantly formula-fed, with breastfed children exhibiting significantly higher IgA targeting of *E. coli* at 3 months of age and significantly lower IgA targeting of *R. gnavus* during the latter half of the first year (Extended Data Fig. 7).

To quantify the stage of development of gut mucosal IgA responses, we randomly selected 20 unrelated children from the healthy twin cohort and generated an RF model based on the IgA indices to the 30 taxa identified in Fig. 1a. The model was then applied to unrelated children represented in the remaining 20 twin pairs ('test set',  $n = 40$ ). Even though the data set was smaller than the one used to generate the

RF model of gut microbiota development, the effort produced a model of development of IgA responses that correlated with donor chronological age (Spearman's correlation for training set and test set, 0.97, and 0.72, respectively; Supplementary Tables 13 and 17).

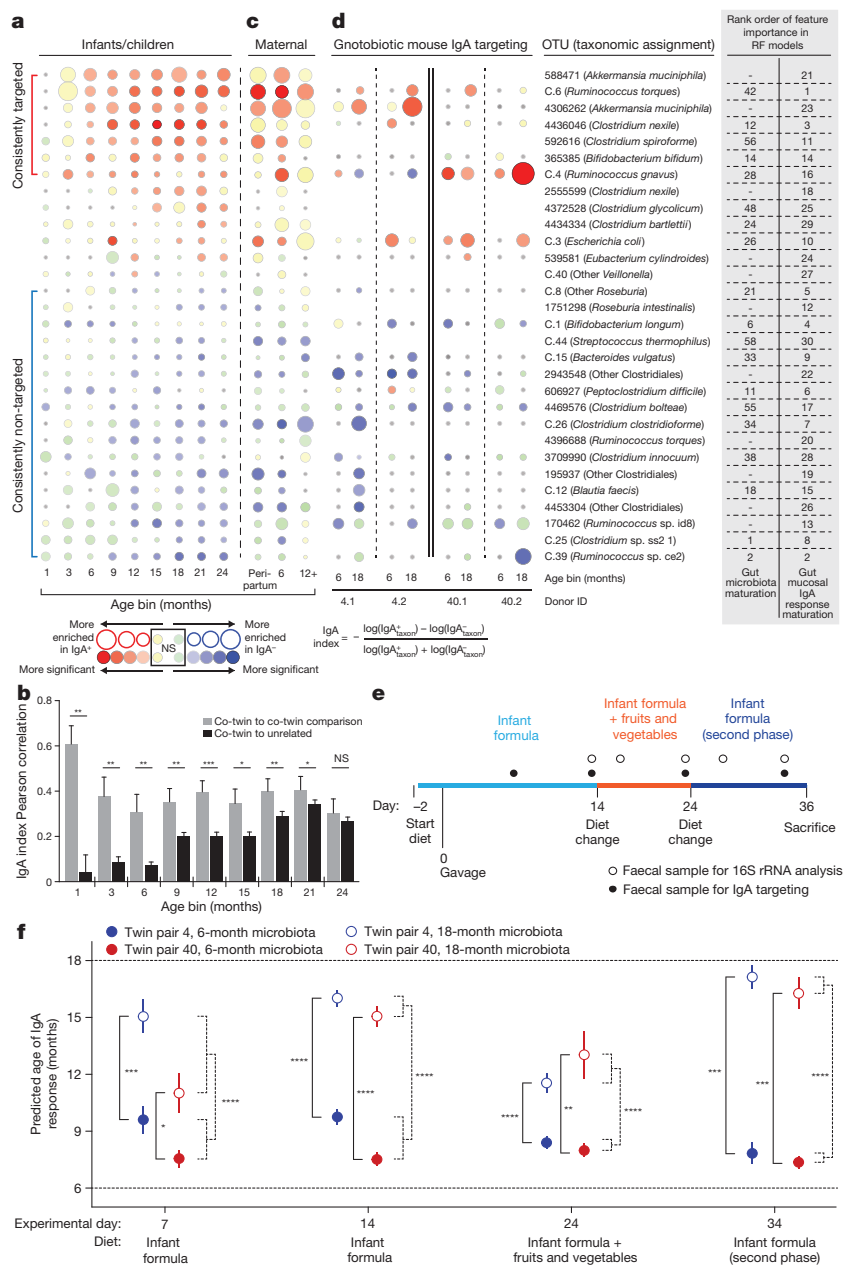
Faecal samples from two twin pairs whose pattern of gut microbiota development was well described by the RF-derived model and whose IgA responses exemplified those of the larger population were selected for transplantation into germ-free mice (pairs 4 and 40 in Supplementary Table 13). Both twin pairs were predominantly formula-fed throughout their first postnatal year (Supplementary Table 2). Faecal samples, collected from each of these four individuals when they were 6 and 18 months old, were introduced into separate groups of male 5-week-old C57BL/6J germ-free mice ( $n = 4$  or 5 mice per donor sample; eight treatment groups). Two days before gavage, all mice were switched from a standard chow diet low in fat and rich in plant polysaccharides to a sterilized human infant formula diet (see Methods). After gavage, animals were maintained on this diet for 14 days and then switched to a diet constructed on the basis of a survey of fruits and vegetables most commonly consumed by infants transitioning to complementary foods<sup>19</sup>. This diet consisted of isocaloric amounts of the powdered infant formula diet and a mixture of sweet potatoes, green beans, bananas, and apples. After 10 days, animals were returned to the infant formula diet for another 10 days. Faecal samples were obtained from recipient mice at frequent intervals throughout all diet phases and subjected to 16S rRNA sequencing and/or to BugFACS (Fig. 1d, e).

Indicator species analysis revealed that the abundances of 15 of the top 60 taxa in the RF-derived model of gut microbiota maturation varied significantly in the context of one or the other diets (false discovery rate (FDR)-corrected  $P < 0.05$  and indicator value  $> 0.5$ ; Supplementary Tables 18 and 19). Most of these taxa responded in the same direction (increased or decreased in abundance) to the different diets, independent of the microbiota donor or donor age (Extended Data Fig. 8). For example, the age-discriminatory OTU C.1 (*B. longum*) and OTU 4439469 (a member of the Ruminococcaceae) have highest mean relative abundances during the first 6 months of postnatal life in members of the twin cohort (Extended Data Fig. 2c); these taxa also exhibited significantly greater relative abundance in the faecal microbiota of recipient gnotobiotic mice when they were consuming the infant formula diet. In contrast, *Anaerostipes caccae* (OTU 259772), which peaks in abundance during the latter half of the first postnatal year (the period corresponding to introduction of complementary foods in our twin study; Extended Data Fig. 2c and Extended Data Fig. 4a), was significantly higher in its abundance during the 'formula plus fruits and vegetables' diet phase (Extended Data Fig. 8).

To determine whether age-associated differences in IgA responses to components of the donors' microbiota could be recapitulated in gnotobiotic mice, we subjected their faecal samples collected at 7, 14, 24, and 34 days after gavage to BugFACS (Supplementary Table 20). IgA responses in mice broadly mirrored those of the human donor population; taxa that were consistently not targeted across members of the twin cohort during the first 2 years of postnatal life (for example, *Clostridium clostridioforme* OTU C.26 and *Clostridium bolteae* OTU 4469576) were generally not targeted in mice colonized with the 6- and 18-month microbiota samples from the two twin pairs, while bacteria targeted in mice belonged to the set of taxa that were consistently IgA targeted in infants/children from postnatal months 6–24 (for example, *Ruminococcus torques* (OTU C.6) and *Akkermansia muciniphila* (OTU 4306262)) (Fig. 1d and Supplementary Table 21).

IgA-targeting of five OTUs varied significantly with the diet oscillation, whether judged by a comparison of the first and second or second and third diet phases (FDR-corrected repeated-measures ANOVA): they included OTUs whose IgA targeting increased during the fruits and vegetables phase (4306262 (*A. muciniphila*) and C.39 (*Ruminococcus* sp. ce2)), and those whose targeting decreased (OTUs C.4 (*R. gnavus*), 4469576 (*C. bolteae*), 4453304 (other Clostridiales)) (Supplementary Table 22). IgA responses were most similar in mice





**Figure 1 | BugFACS-based analysis of the development of gut mucosal IgA responses in healthy USA twin pairs.**

**a**, Bubble plots for OTUs significantly enriched in either the IgA<sup>+</sup> or IgA<sup>-</sup> fractions across twin microbiota samples representing at least three age bins ( $n = 534$  faecal samples analysed;  $6.7 \pm 1.7$  (mean  $\pm$  s.d.) samples per individual from the 1-, 3-, 6-, 9-, 12-, 15-, 18-, 21-, and 24-month age bins). The size of the circle indicates the magnitude of enrichment (mean IgA index value) in either fraction. Colour intensity indicates the statistical significance, as judged by a Wilcoxon signed-rank test. Grey indicates that a taxon was not detected in any sample at that time point. Taxa are ordered by their mean IgA indices defined across all time points. The rank order of feature importance for OTUs in the RF-derived models of 'gut microbiota maturation' and 'gut mucosal IgA response maturation' is indicated. The upper bracket to left of the figure encompasses OTUs identified as consistently IgA-targeted after the first three postnatal months, while the lower bracket highlights OTUs designated as consistently not targeted in multiple age bins after the first 3 postnatal months. **b**, Pearson's correlation coefficients were used to compare the similarity of IgA index profiles of twins and unrelated children within a given age bin ( $n = 9$  total age bins). The correlation coefficient between members of a twin pair in a given age bin was compared with the mean of the correlations between that pair and all other age-matched unrelated children in the cohort (mean  $\pm$  s.e.m.). Statistical significance was defined with the paired Wilcoxon test. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ . **c**, Samples from mothers of the twin pairs were collected during the peripartum period, at 6 months after birth, and at  $\geq 12$  months after birth. Samples ( $n = 106$ ) were subjected to BugFACS and analysed as above. **d**, Modelling development of gut IgA responses in young gnotobiotic mice colonized with the faecal microbiota from two twin pairs and fed infant formula and complementary foods. Bubble plots for OTUs shown in **a** that were present in BugFACS fractions prepared from the faecal microbiota of mice harbouring transplanted faecal microbiota from 6- and 18-month-old co-twins belonging to the two twin pairs ( $n = 4$  or 5 mice per group, total of 143 faecal samples analysed). Data were pooled from the four indicated time points in the diet oscillation. **e**, Experimental design for gnotobiotic mouse experiment. **f**, 'Age' of the IgA responses documented in mice as a function of donor microbiota and diet. The RF model of human gut mucosal IgA response maturation was applied to BugFACS data sets generated from mouse samples. Data are pooled within recipients of twin pair microbiota from a given age bin for each time point ( $n = 4$  or 5 mice per group, 143 faecal samples analysed). Mean values  $\pm$  s.e.m. are plotted. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; \*\*\*\* $P < 0.0001$  (Mann-Whitney  $U$ -test for the indicated comparisons).

harbouring a given donor microbiota, and were more similar within members of a twin pair than between unrelated children (Extended Data Fig. 9).

Applying our RF-derived model of maturation of human gut mucosal IgA responses to the mouse BugFACS data set showed that animals recapitulated distinct age-associated differences in mucosal IgA responses to the (transplanted) human microbiota; that is, for both twin pairs, the state of maturation of the IgA response in mice was significantly greater when animals were colonized with the 18-month compared with the 6-month donors' communities. Remarkably, this significant difference in age-associated responses for a given co-twin or twin pair microbiota was evident in both diet contexts (Fig. 1f and Supplementary Table 21). We concluded that IgA responses to members of the 6-month-old gut microbiota were shared across the two twin pairs and robust to the transition to complementary foods. The fact that distinctive responses to the 18-month compared with the 6-month

microbiota were identified in recipient mice, even in the context of a milk (formula) diet, supports the notion that 'intrinsic' properties of community members (for example, properties not clearly related to taxonomy or obviously affected by community composition) play a dominant role in dictating the gut mucosal IgA targeting response.

Our findings point to several directions for future investigation. The stability of the IgA molecule, the ease and safety of obtaining faecal samples, and the ability to sort members of a faecal microbiota sample into IgA-enriched versus non-enriched fractions provide a way to non-invasively quantify states of development of the gut mucosal immune system as a function of different host and environmental factors. BugFACS offers an opportunity to identify previously unappreciated 'IgA deficiencies' presenting not as a lack of, or reduction in, the amount of total IgA in the gut lumen, but rather as aberrant patterns of IgA targeting. The effects of such deficiencies would need to be examined with the understanding that barrier function can be affected

by multiple factors besides IgA, including for example mucin<sup>20</sup>, and less well understood elements<sup>21</sup>. BugFACS also provides a way to explore how IgA targeting of bacterial taxa varies as a function of their proximity to the intestinal epithelium and their location along the length of the gut. The importance of the small intestine as a source of the T-cell-independent IgA response to members of the microbiota was highlighted in a recent study<sup>4</sup>.

In principle, deviations from the pattern of convergence of IgA responses observed in the present study could occur in scenarios where colonization is abnormal, resulting in pathological immune responses in anatomically distant locations, such as that observed in asthma<sup>22</sup> or various autoimmune/immunoinflammatory disorders. One obvious next step is to assess the generalizability of the shared pattern observed in this study by characterizing healthy members of birth cohorts representing different geographical areas, distinctive cultural and dietary traditions, and living environments with varying degrees of sanitation. Gut microbiota development is impaired in children with undernutrition<sup>12</sup>. Given that undernutrition is associated with impaired gut barrier function and responses to particular vaccines<sup>23</sup>, a comparison of the development of gut mucosal IgA responses to members of the microbiota in healthy and undernourished members of birth cohorts could provide a metric for disease classification, assessment of the impact of enteropathogen infection/burden, and a means of assessing the efficacy of current or new therapeutic interventions, including approaches for oral vaccination.

The ability to re-enact and recapitulate features of the development of gut mucosal IgA responses to human donor gut microbial communities in wild-type or genetically manipulated gnotobiotic mice should help delineate the mechanisms that control the temporal evolution and specificity of IgA responses to members of the gut community, the effects of the IgA response on targeted microbes and other members of the microbiota, as well as the impact on host biology. As such, these models could be used to identify new strategies for deliberately manipulating mucosal barrier/immune function, including food-based and/or microbial interventions.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 5 November 2015; accepted 21 March 2016.

Published online 25 May 2016.

1. Suzuki, K. *et al.* Aberrant expansion of segmented filamentous bacteria in IgA-deficient gut. *Proc. Natl Acad. Sci. USA* **101**, 1981–1986 (2004).
2. Johansen, F. E. *et al.* Absence of epithelial immunoglobulin A transport, with increased mucosal leakiness, in polymeric immunoglobulin receptor/secretory component-deficient mice. *J. Exp. Med.* **190**, 915–922 (1999).
3. Peterson, D. A., McNulty, N. P., Guruge, J. L. & Gordon, J. I. IgA response to symbiotic bacteria as a mediator of gut homeostasis. *Cell Host Microbe* **2**, 328–339 (2007).
4. Bunker, J. J. *et al.* Innate and adaptive humoral responses coat distinct commensal bacteria with immunoglobulin A. *Immunity* **43**, 541–553 (2015).
5. Macpherson, A. J. *et al.* A primitive T cell-independent mechanism of intestinal mucosal IgA responses to commensal bacteria. *Science* **288**, 2222–2226 (2000).
6. Mathias, A. & Corthésy, B. N-Glycans on secretory component: mediators of the interaction between secretory IgA and Gram-positive commensals sustaining intestinal homeostasis. *Gut Microbes* **2**, 287–293 (2011).

7. Palm, N. W. *et al.* Immunoglobulin A coating identifies colitogenic bacteria in inflammatory bowel disease. *Cell* **158**, 1000–1010 (2014).
8. Kau, A. L. *et al.* Functional characterization of IgA-targeted bacterial taxa from undernourished Malawian children that produce diet-dependent enteropathy. *Sci. Transl. Med.* **7**, 276r24 (2015).
9. Brandtzaeg, P. The mucosal immune system and its integration with the mammary glands. *J. Pediatr.* **156** (Suppl.), S8–S15 (2010).
10. Gustafson, C. E. *et al.* Limited expression of APRIL and its receptors prior to intestinal IgA plasma cell development during human infancy. *Mucosal Immunol.* **7**, 467–477 (2014).
11. Rogosch, T. *et al.* IgA response in preterm neonates shows little evidence of antigen-driven selection. *J. Immunol.* **189**, 5449–5456 (2012).
12. Subramanian, S. *et al.* Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature* **510**, 417–421 (2014).
13. Blanton, L. V. *et al.* Gut bacteria that prevent growth impairments transmitted by microbiota from malnourished children. *Science* **351**, aad3311 (2016).
14. Goodrich, J. K. *et al.* Human genetics shape the gut microbiome. *Cell* **159**, 789–799 (2014).
15. Hansen, E. E. *et al.* Pan-genome of the dominant human gut-associated archaeon, *Methanobrevibacter smithii*, studied in twins. *Proc. Natl Acad. Sci. USA* **108** (Suppl. 1), 4599–4606 (2011).
16. Dominguez-Bello, M. G. *et al.* Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc. Natl Acad. Sci. USA* **107**, 11971–11975 (2010).
17. Sela, D. A. & Mills, D. A. The marriage of nutrigenomics with the microbiome: the case of infant-associated bifidobacteria and milk. *Am. J. Clin. Nutr.* **99**, 697S–703S (2014).
18. Dufrene, M. & Legendre, P. Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecol. Monogr.* **67**, 345–366 (1999).
19. Siega-Riz, A. M. *et al.* Food consumption patterns of infants and toddlers: where are we now? *J. Am. Diet. Assoc.* **110** (Suppl.), S38–S51 (2010).
20. Rogier, E. W., Frantz, A. L., Bruno, M. E. & Kaetzel, C. S. Secretory IgA is concentrated in the outer layer of the colonic mucus along with gut bacteria. *Pathogens* **3**, 390–403 (2014).
21. Lyte, M., Vulchanova, L. & Brown, D. R. Stress at the intestinal surface: catecholamines and mucosa-bacteria interactions. *Cell Tissue Res.* **343**, 23–32 (2011).
22. Arrieta, M. C. *et al.* Early infancy microbial and metabolic alterations affect risk of childhood asthma. *Sci. Transl. Med.* **7**, 307ra152 (2015).
23. Gilmartin, A. A. & Petri, W. A., Jr. Exploring the role of environmental enteropathy in malnutrition, infant development and oral vaccine response. *Phil. Trans. R. Soc. B* **370**, 20140143 (2015).

Supplementary Information is available in the online version of the paper.

**Acknowledgements** We thank D. O'Donnell, M. Karlsson, J. Serugo, and S. Wagoner for help with gnotobiotic husbandry; S. Deng, J. Guruge, J. Hoisington-Lopez and M. Meier for technical assistance; G. Dantas for help with maintaining our archive of de-identified human samples; and N. Griffin for comments about facets of the data analysis. This work was supported by grants from the National Institutes of Health (DK30292, DK052574), the Children's Discovery Institute, the Bill & Melinda Gates Foundation, and the Crohn's and Colitis Foundation of America. J.D.P. is a member of the Washington University Medical Scientist Training Program (National Institutes of Health GM007200).

**Author Contributions** B.B.W., P.I.T., M.I. and G.D. designed, enrolled and collected specimens from participants in the twin study. J.D.P. performed BugFACS and 16S rRNA analyses on human faecal samples. J.D.P. and J.I.G. designed the gnotobiotic mouse experiments; J.D.P. and Y.P. performed these experiments. J.D.P., A.L.K., L.V.B., Y.P., and J.I.G. analysed the data. J.D.P. and J.I.G. wrote the paper.

**Author Information** 16S rRNA sequences in raw format before post-processing and data analysis have been deposited at the European Nucleotide Archive under project PRJEB11697. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.I.G. ([jgordon@wustl.edu](mailto:jgordon@wustl.edu)).

## METHODS

**Human studies.** Protocols used for recruitment of participants, obtaining informed consent, collecting and de-identifying faecal samples, and acquiring and de-identifying clinical metadata were all approved by the Human Research Protection Office of Washington University School of Medicine. A total of 40 twin pairs were included in this study; this number was not determined by a power calculation.

All breast milk provided was from the mothers of the twins themselves and was not pasteurized. Breast milk was given from the breast directly or from bottles after being expressed by the mother. Expressed milk was given immediately or stored temporarily in home freezers; in the latter case, mothers were instructed to thaw their milk in warm water.

**Determination of zygosity.** Zygosity testing of same-gender twins was performed on residual blood samples obtained for clinical care, or samples obtained at the time of mandatory Missouri-state metabolic testing. Short tandem repeat polymorphic DNA markers were amplified from blood DNA by PCR, labelled with fluorescent markers, and separated by capillary electrophoresis to distinguish different alleles at each of ten different loci (D3S1358, vWA, FGA, Amelogenin, D8S1179, D21S11, D18S51, D5S818, D13S317, and D7S820).

**V4-16S rRNA gene sequencing and data analysis.** Faecal samples were quickly frozen at  $-20^{\circ}\text{C}$  and subsequently stored at  $-80^{\circ}\text{C}$ . Samples were pulverized in liquid nitrogen and DNA was extracted from an aliquot of the material ( $130 \pm 36$  mg; mean  $\pm$  s.d.) by bead-beating in a solution consisting of 500  $\mu\text{l}$  of phenol:chloroform:isoamyl alcohol (25:24:1), 210  $\mu\text{l}$  of 20% SDS, and 500  $\mu\text{l}$  of buffer A (200 mM NaCl, 200 mM Trizma base, 20 mM EDTA). DNA was further purified with Qiaquick columns (Qiagen), eluted in 70  $\mu\text{l}$  of Tris-EDTA (TE) buffer, and quantified (Quant-iT dsDNA broad range kit; Invitrogen). The concentration of each DNA sample was normalized to 1 ng  $\mu\text{l}^{-1}$  and the DNA was subjected to PCR with phased, barcoded primers directed against variable region 4 of the bacterial 16S rRNA gene<sup>24</sup>. Amplicons were quantified as above, pooled, and sequenced on an Illumina MiSeq instrument (paired-end 250 nucleotide reads). Paired-end reads were merged (FLASH, version 1.2.6). De-multiplexed reads were clustered into OTUs with the 97% identity sequence set from the GreenGenes 2013 reference database and QIIME version 1.8 (ref. 25). An 'abundance-filtered data set' was generated by selecting OTUs that were detected at  $>0.1\%$  relative abundance in  $\geq 1\%$  of the samples; only these OTUs were considered for further analysis.

Taxonomy was assigned to 97% nucleotide sequence identity OTUs with RDP 2.4, as described previously<sup>8</sup>. Taxonomically related OTUs that shared a high degree of rank order co-linearity of their abundance (Spearman's  $\rho > 0.7$ ) were consolidated as follows: (1) for each family-level taxon that was detected in the abundance-filtered data set, a list of OTUs belonging to the family was generated; (2) the abundances of OTUs within this family across all samples analysed were then correlated with each other to generate Spearman's correlation coefficients for each OTU–OTU comparison; and (3) counts for OTUs that shared a high degree of co-linearity with each other were then combined to generate the consolidated OTUs according to the scheme that is illustrated in Extended Data Fig. 1 (see Supplementary Table 5 for a list of OTUs that were consolidated; note that OTUs used to generate a 'consolidated OTU' shared on average  $99.3 \pm 0.4\%$  (mean  $\pm$  s.d.) nucleotide sequence identity in their V4-16S rRNA sequences; OTUs that did not satisfy the threshold cutoff for co-linearity in abundance were not consolidated). A new OTU table was then generated, consisting of consolidated OTUs that were assigned a new OTU sequence identity and all other non-consolidated OTUs. These OTU tables were then rarefied to depths of 1,000 reads for the BugFACS-related analyses described below, and 5,000 reads for all other analyses.

**BugFACS of human samples.** A separate aliquot of a pulverized frozen faecal sample was transferred to a pre-weighed 1.5 ml microcentrifuge tube (Life Technologies), and processed as described previously with some minor modifications<sup>8</sup>. Samples were resuspended in 1 ml of PBS, vortexed at room temperature for 5 min (1,500 rotations per minute), and then placed on ice for 5 min to allow large particulate matter to settle by gravity. A volume equivalent to 5 mg of pulverized faecal material was passed through a nylon 70  $\mu\text{m}$  mesh filter (BD). One millilitre of ice cold PBS was added to each filtered sample, which was then centrifuged at 10,000g for 3 min ( $4^{\circ}\text{C}$ ). The resulting supernatant was discarded and the cell pellet was resuspended in 100  $\mu\text{l}$  of a 1:50 dilution of goat anti-human IgA conjugated to DyLight 650 (Abcam; catalogue number ab96998). Samples were subsequently incubated on ice in the dark for 30 min, washed with 1 ml of PBS, and resuspended in a 1:4,000 dilution of SytoBC bacterial DNA stain (Life Technologies; prepared in HEPES-NaCl buffer (0.9% NaCl, 10 mM HEPES)), immediately before introduction into a FACSaria III cell sorter (Becton Dickinson).

For each sample, 50,000 cytometer 'events' were recovered from the 'Input', 'IgA<sup>+</sup>', and 'IgA<sup>-</sup>' gates (for details of the sorting protocol and gating strategies, see ref. 8). Additionally, samples of sheath fluid were collected immediately before and

after sorting to allow assessment of any potential contaminants in fluid lines. Sorted fractions and control sheath fluid samples were frozen and stored at  $-20^{\circ}\text{C}$ . Each BugFACS-sorted fraction was subjected to V4-16S rRNA gene PCR in triplicate 20  $\mu\text{l}$  reactions. Each reaction contained 2  $\mu\text{l}$  of 10X HiFi PCR Buffer (Invitrogen), 0.8  $\mu\text{l}$  of 50 mM magnesium sulfate (Invitrogen), 0.4  $\mu\text{l}$  of dNTP mix (Invitrogen), 0.16  $\mu\text{l}$  of Platinum Taq (Invitrogen), 1  $\mu\text{l}$  of a 5  $\mu\text{M}$  stock of forward PCR primer, 1  $\mu\text{l}$  of 5  $\mu\text{M}$  barcoded reverse PCR primer<sup>26</sup>, 2.5  $\mu\text{l}$  of BugFACS sorted cells, and 12.1  $\mu\text{l}$  of water. A negative control reaction with no sorted cells was included for each barcoded primer. The following PCR conditions were used:  $95^{\circ}\text{C}$  for 10 min followed by 31 cycles of  $95^{\circ}\text{C}$  for 30 s,  $53^{\circ}\text{C}$  for 30 s, and  $68^{\circ}\text{C}$  for 45 s, followed by  $68^{\circ}\text{C}$  for 2 min. Triplicate reactions were pooled and subjected to 1% agarose gel electrophoresis to verify the presence of a PCR product (these gels also contained negative control reactions). If any of the three sorted fractions from a given sample failed to amplify successfully, PCRs were repeated for 34 cycles for all three fractions under the same temperature cycling conditions. PCR-amplified fractions were pooled in equal proportion. Although amplicon bands were not visible for sheath fluid controls, a set volume of these reactions was also included in the sequencing pool. Pooled amplicons were purified with magnetic beads (AMPure XP, Agencourt) and subjected to multiplex sequencing (paired-end 250 nucleotide reads) on a MiSeq instrument as above.

After OTU picking, but before abundance filtering, sheath fluid-contaminating OTUs were identified as sequences that constituted  $>1\%$  of the reads in both the pre- and post-sort sheath fluid samples for a given day. Contaminants that were identified on more than 2 days were removed from the OTU table. If multiple genera within a family-level taxon were identified as sheath contaminants, the entire family was removed from the OTU table. This list included the following families: Burkholderiaceae, Xanthomonadaceae, Comamonadaceae, Brucellaceae, Pseudomonadaceae, Xanthobacteraceae, and Alcaligenaceae. Altogether, OTUs belonging to these families accounted for less than 0.05% of all sequences in the twin pair, maternal, and mouse faecal samples, and for less than 2% of all sequences in samples subjected to BugFACS. IgA indices were subsequently calculated for a given taxon in a given sample if that taxon comprised  $\geq 0.5\%$  of the 16S rRNA reads in either the IgA<sup>+</sup> or IgA<sup>-</sup> fraction.

**RF modelling.** RF modelling of gut microbiota development was performed with the 'randomForest' package<sup>27</sup> in R. Input data consisted of OTU data rarefied to a depth of 5,000 V4-16S rRNA reads per faecal sample. Feature importance scores for each OTU in the data set were calculated by randomly selecting one co-twin from half of the twin pairs ( $n = 20$  individuals). An RF model was generated from this subset of data. Randomization and this process of model construction were performed 100 times (100 trees per model). Feature importance scores were extracted from each model, averaged across the 100 models, and used to rank the OTUs from highest to lowest feature importance.

To estimate the number of OTUs needed to build a sparse model, a new set of RF models was generated by selecting one co-twin from half of the twin pairs in the cohort as above, and evaluating the performance of the model (Spearman's  $\rho$  and the adjusted  $r^2$  of a linear model as metrics) when applied to (1) the individuals used to generate the model, (2) their co-twins, and (3) all unrelated faecal samples ('training', 'co-twin', and 'test' sets, respectively). A series of models was built with increasing numbers of OTUs, starting with the OTU assigned the highest feature importance score, and sequentially adding OTUs in decreasing order of feature importance. For each model of a different size, ten randomizations were performed; performance of the model was averaged across the independent replicates to generate standard error measurements. The subset of 25 OTUs with highest rank order of feature importance scores was used to create a sparse model. This sparse model, generated from samples collected during the first 36 months of postnatal life, was applied to 16S rRNA data sets generated from faecal samples collected between 1 and 24 months of age to predict chronological age in members of the 'training', 'co-twin', and 'test' subsets as described above. A parallel RF-derived model was generated from IgA index data for the 30 OTUs shown in Fig. 1a. If a taxon was not detected in either the IgA<sup>+</sup> or IgA<sup>-</sup> fraction, it was given a value of 0 before model construction. This model was applied to the 'training', 'co-twin', and 'test' sets.

OTUs were reassigned to incorporate data sets from all three countries (USA, Bangladesh, and Malawi), resulting in a second set of consolidated OTUs (see Supplementary Table 9). Feature importance scores were calculated by iteratively regressing each country's training set of samples 100 times against chronological age (100 trees per model); OTUs were ranked by the mean values of their feature importance scores across the 100 models. The 25 most age-discriminatory OTUs were used to generate each respective country's sparse RF model. Each model was used to predict the microbiota ages of members of that country's corresponding test set, as well as the microbiota ages of all members of the healthy cohorts from each of the other two countries. Spearman's correlation coefficients were generated



by building each sparse RF model ten times, correlating predicted microbiota ages with chronological ages, and averaging the coefficients.

**Animal studies.** All experiments involving mice were performed according to protocols that were in compliance with ethical regulations and approved by the Washington University Animal Studies Committee. No inclusion or exclusion criteria were established; all animals studied were included in our analyses.

**Gnotobiotic mouse husbandry.** Germ-free 5-week-old male C57BL/6J mice (*Mus musculus*) were maintained on a strict 12 h light cycle (lights on at 6:00) in flexible plastic gnotobiotic isolators (Class Biologically Clean). Mice were weaned onto an autoclaved, standard mouse chow diet low in fat and rich in plant polysaccharides (B&K Universal; diet 7378000). Two days before introduction of human donor faecal samples by gavage, 5-week-old animals were switched to the human infant formula diet.

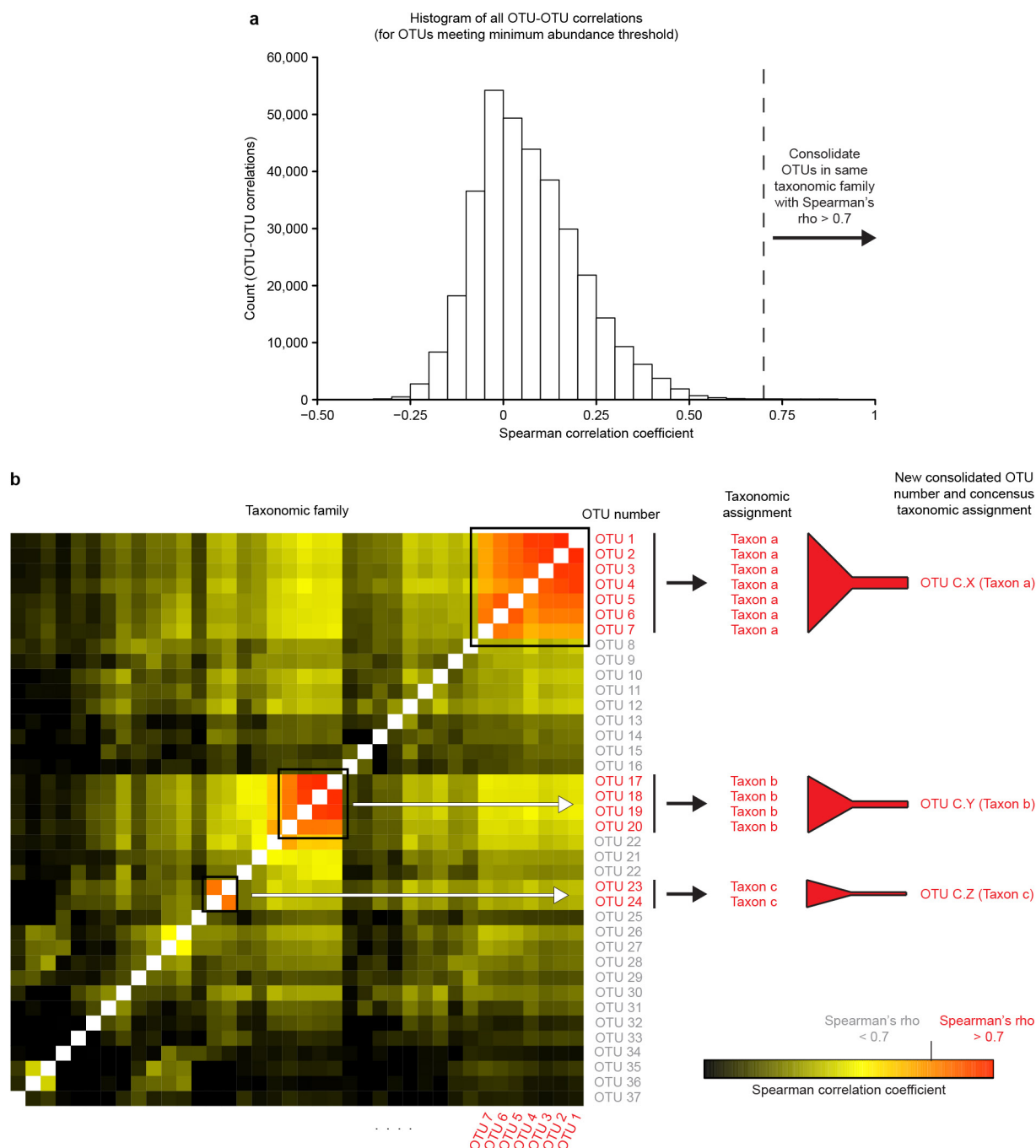
**Diets.** The infant formula diet consisted of a mixture of Similac 'Sensitive with Iron' infant formula and unflavoured whey protein powder (GNC) mixed at a ratio of 11:1 (w/w). This powdered diet was reconstituted in the gnotobiotic isolator on a daily basis with sterile water. The infant formula plus fruit and vegetable diet was based on a survey of the fruits and vegetables most commonly consumed by infants transitioning to complementary foods<sup>19</sup>, and consisted of isocaloric amounts of the powdered infant formula diet and a mixture of 1:1:1:1 ratio (by mass) of sweet potatoes, green beans, bananas, and apples (Gerber 1st Foods). Formula was irradiated as a powder. Fruits and vegetables were irradiated in their original plastic containers before the start of the experiment (25–30 Gy; Steris Isomedix) and mixed with the irradiated formula powder. When mice were consuming infant formula diet alone, fresh food was prepared daily within the gnotobiotic isolator and presented to animals in sterile plastic trays that were changed daily. When animals were given the mixture of formula and fruits/vegetables, food was prepared every other day, and new aliquots given to animals in fresh trays daily. Bedding was changed with each phase of the diet oscillation; within a given diet phase, bedding was changed every 2–3 days.

**Microbiota transplants.** A given pulverized frozen human faecal sample ( $353 \pm 184$  mg; mean  $\pm$  s.d.) was transferred to an anaerobic Coy chamber (atmosphere 75% N<sub>2</sub>, 20% CO<sub>2</sub>, 5% H<sub>2</sub>) in a 2 ml Axygen screw topped tube. The tube was then opened and its contents were transferred to a 50 ml conical shaped polypropylene tube (Falcon). The faecal material was suspended in 10 ml of sterile PBS supplemented with 0.1% L-cysteine (Sigma) by vortexing with sterile glass beads (2 mm in diameter). The suspension was passed through a nylon 100  $\mu$ m mesh filter (BD) and the filtrate was mixed with an equal volume of 30% glycerol in PBS/0.1% cysteine. Aliquots (1.2 ml) of this suspension were placed in amber glass vials, each of which was sealed with a crimp top, and frozen at  $-80^{\circ}\text{C}$ . Tubes were thawed, and transferred into gnotobiotic isolators (with surface sterilization achieved by treatment with Clidox). Aliquots (200  $\mu$ l) were then introduced into each germ-free mouse in a given experimental group by oral gavage. A total of 38 animals were used for this study ( $n = 4$  or 5 mice per donor microbiota). This size of each treatment group was not based on a formal power calculation but was informed by our previous work described in ref. 8. There was no randomization of mice for this study; male C57BL/6J animals in each group were age- and weight-matched before gavage. Investigators were not blinded to the donor microbiota.

**BugFACS of mouse faecal samples.** The protocol used was similar to that described above for human faecal samples with several modifications. Faecal pellets were resuspended in PBS, vortexed, and a volume equivalent to 5 mg of faecal material was passed through a nylon 70  $\mu$ m mesh filter. After washing with PBS, cells were incubated for 30 min on ice in the dark with a polyclonal goat antibody directed against mouse IgA conjugated to DyLight 650 (Abcam; catalogue number ab97014; diluted 1/50 in PBS/0.5% (w/v) bovine serum albumin). On each day that BugFACS was performed, a positive control of pooled material from all mouse faecal samples analysed on that day and stained with anti-mouse IgA antibody was used to verify staining, while a negative control of the same pooled faecal material stained with the anti-human IgA antibody (conjugated to DyLight 650, see above) was used as an isotype control.

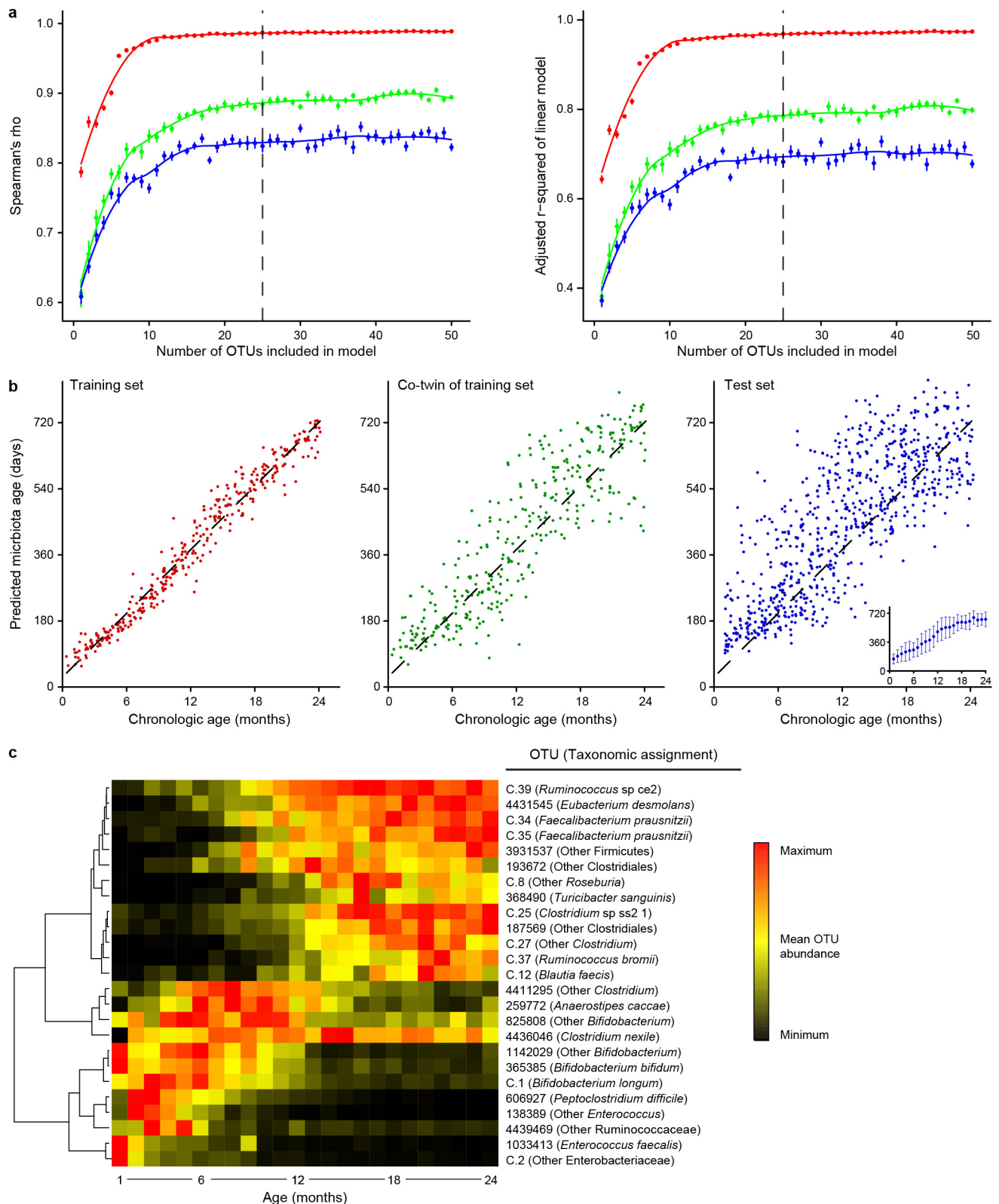
**Statistics.** Statistical analyses, RF modelling, generation of plots, OTU consolidation, and OTU table rarefaction were performed in the R programming environment (R version 3.1.1) or Prism 6.0. For presentations of data in which group means are compared, confidence in mean values is displayed as the s.e.m. Mann–Whitney *U*-tests and Student's *t*-tests were all two-tailed. FDR correction of *P* values was performed with the Benjamini–Hochberg procedure. Indicator species analysis was performed with the 'indicspecies' package in R<sup>28</sup>. PERMANOVA tests were performed with the 'vegan' package in R<sup>29</sup>. For PERMANOVA of IgA indices, the matrix of Pearson's product moment correlation coefficients was converted to a dissimilarity matrix with the formula  $X_{\text{dissimilarity}} = (1 - X_{\text{similarity}})/2$ , where *X* represents a given sample-to-sample comparison. We performed two separate analyses: in the first, only the effects of zygosity, delivery mode, age bin, and twin pair were considered; a second analysis was performed on the subset of samples for which feeding data were available to evaluate the effects of milk feeding practices, with age bin and twin pair included as covariates. In both cases, 999 permutations were performed. Linear mixed-effects modelling was performed with the 'lmerTest' package in R<sup>30</sup>. Sex, delivery mode, zygosity, and feeding predominance were tested as fixed effects, with age bin, twin pair, and the infant/child study identifier treated as nested random effects. Similar results were obtained with either the Satterthwaite or the Kenward–Roger approximation for denominator degrees of freedom.

24. Faith, J. J., Ahern, P. P., Ridaura, V. K., Cheng, J. & Gordon, J. I. Identifying gut microbe-host phenotype relationships using combinatorial communities in gnotobiotic mice. *Sci. Transl. Med.* **6**, 220ra11 (2014).
25. Rideout, J. R. *et al.* Subsampled open-reference clustering creates consistent, comprehensive OTU definitions and scales to billions of sequences. *PeerJ* **2**, e545 (2014).
26. Caporaso, J. G. *et al.* Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl Acad. Sci. USA* **108** (Suppl. 1), 4516–4522 (2011).
27. Liaw, A. & Weiner, M. Classification and regression by randomForest. *R News* **2** (3), 18–22 (2002).
28. De Caceres, M. & Legendre, P. Associations between species and groups of sites: indices and statistical inference. *Ecology* **90**, 3566–3574 (2009).
29. Oksanen, J. *et al.* vegan: Community Ecology Package. <http://CRAN.R-project.org/package=vegan> (2014).
30. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest: tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). <https://cran.r-project.org/package=lmerTest> (2013).



**Extended Data Figure 1 | Method used for OTU consolidation.** **a**, OTU consolidation was performed to limit pseudo-duplication of taxa. 'Counts' on the  $y$  axis refer to the number of OTU-OTU correlations falling within a given range of Spearman's correlation values shown on the  $x$  axis ( $n = 341,640$  OTU-OTU comparisons; see Methods for details). **b**, A subset of the matrix used to derive the distribution shown in **a** illustrates how OTUs within a single family-level taxon are consolidated. In this

example, three clusters composed of OTUs with Spearman's correlation coefficients of  $> 0.7$  are identified and the abundances of their constituent OTUs are summed. Each OTU cluster is assigned an identifier number with the prefix 'C.' and given a consensus taxonomic assignment (see Supplementary Table 5). Note that the OTUs used to generate a given 'consolidated OTU' shared  $99.3 \pm 0.4\%$  (mean  $\pm$  s.d.) nucleotide sequence identity in their V4-16S rRNA nucleotide sequences.

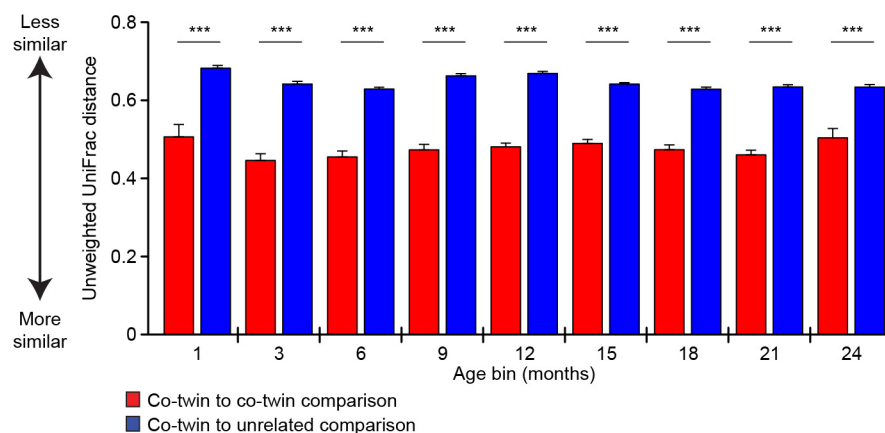


**Extended Data Figure 2 | Modelling development of the gut microbiota during the first 24 months of life in healthy twins.** **a**, To estimate the number of OTUs needed to maximize predictive accuracy, OTUs were iteratively added to a series of RF models, starting with the OTU with the highest feature importance score and adding additional OTUs in order of decreasing feature importance. To evaluate performance of the model, members of the 40 twin cohort were randomly assigned to 'training', 'co-twin of training', and 'test' sets (red, green, and blue, respectively) ten times, and the Spearman's correlation coefficient and adjusted  $r^2$  of a linear model were calculated for a given model size ( $n = 10$  models for each data

point, mean  $\pm$  s.e.m. values are plotted). The dashed vertical line indicates performance of a 25 OTU model across the three different sets.

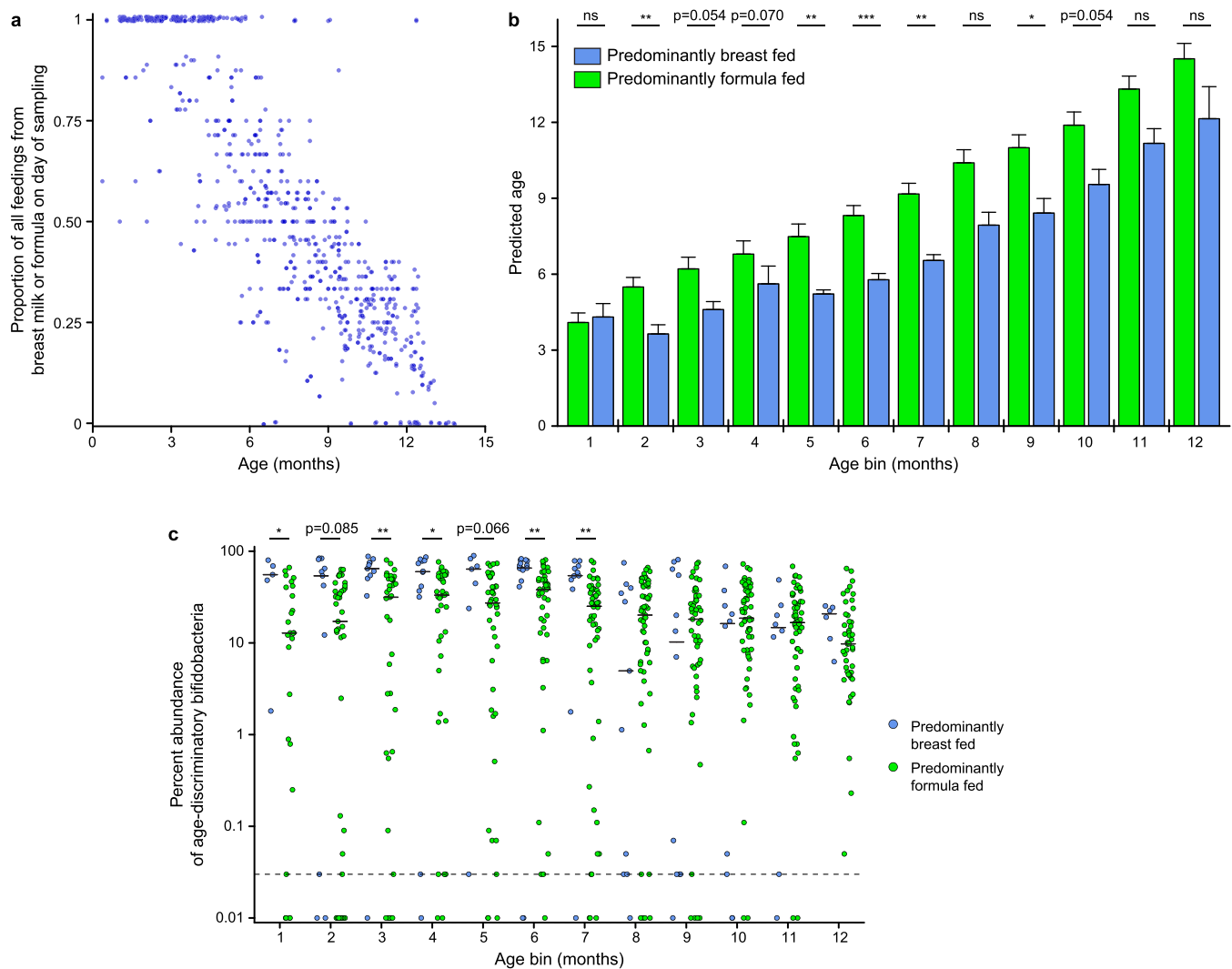
**b**, Predicted age was calculated for all faecal microbiota samples with a sparse 25 OTU RF-generated model. Chronological versus model-predicted age is plotted for each of the three data subsets ( $n = 1,477$  faecal samples). The inset shows mean  $\pm$  s.d. values for predicted microbiota age of samples in each monthly age bin. **c**, Heatmap of mean abundances over the first 24 months of life for the 25 OTUs used to generate the sparse model. Taxa are normalized by row, with hierarchical clustering (complete linkage;  $n = 1,477$  faecal samples).





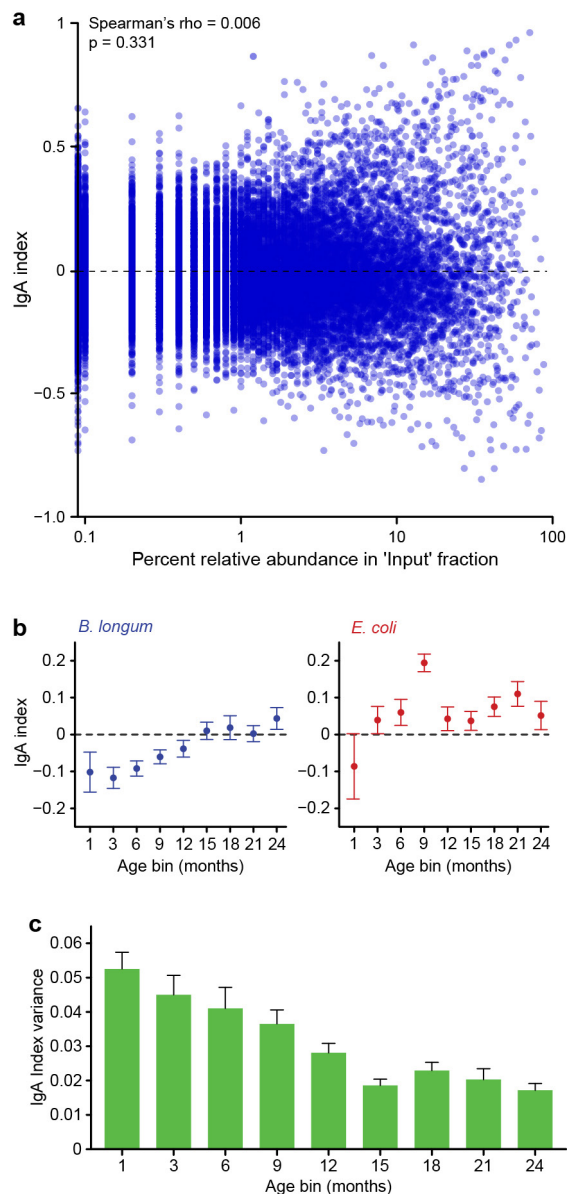
**Extended Data Figure 3 | Similarity of faecal microbiota composition within and between twin pairs.** Similarity in composition of the faecal microbiota within and between twin pairs was analysed with unweighted UniFrac distance calculated before OTU consolidation. Statistical significance was evaluated with the paired Wilcoxon test for twin–twin versus twin–unrelated comparisons. Mean values + s.e.m. are plotted

( $n = 205$  paired comparisons). \*\*\* $P < 0.001$ . The results indicate that the overall phylogenetic composition of the faecal microbiota is more similar in infants/children sharing a common living environment and genetic background than between unrelated individuals; this is apparent as early as the first month of life and does not change significantly over the ensuing 23 months.



**Extended Data Figure 4 | Feeding status and microbiota composition during the first year of life.** **a**, The proportion of all feedings on the day of faecal sampling that consisted of *either* formula or breast milk ( $n = 746$  observations). **b**, Microbiota age, defined by the sparse RF-derived model, compared for participants that were predominantly breast fed ( $\geq 50\%$  of milk feeding) or predominantly formula fed across different age bins. Mean values  $\pm$  s.e.m. are plotted ( $n = 681$  faecal samples). \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$  (Mann–Whitney  $U$ -test). **c**, Aggregate percentage

relative abundance of age-discriminatory bifidobacteria included in the sparse RF-derived model; differences in their representation in the faecal microbiota as a function of breast or formula feeding evaluated in each age bin are shown. Horizontal lines within each column represent the median values; the horizontal dashed line represents the lower limit of detection. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$  (Mann–Whitney  $U$ -test comparing samples obtained from breast versus formula fed individuals;  $n = 681$  faecal samples).

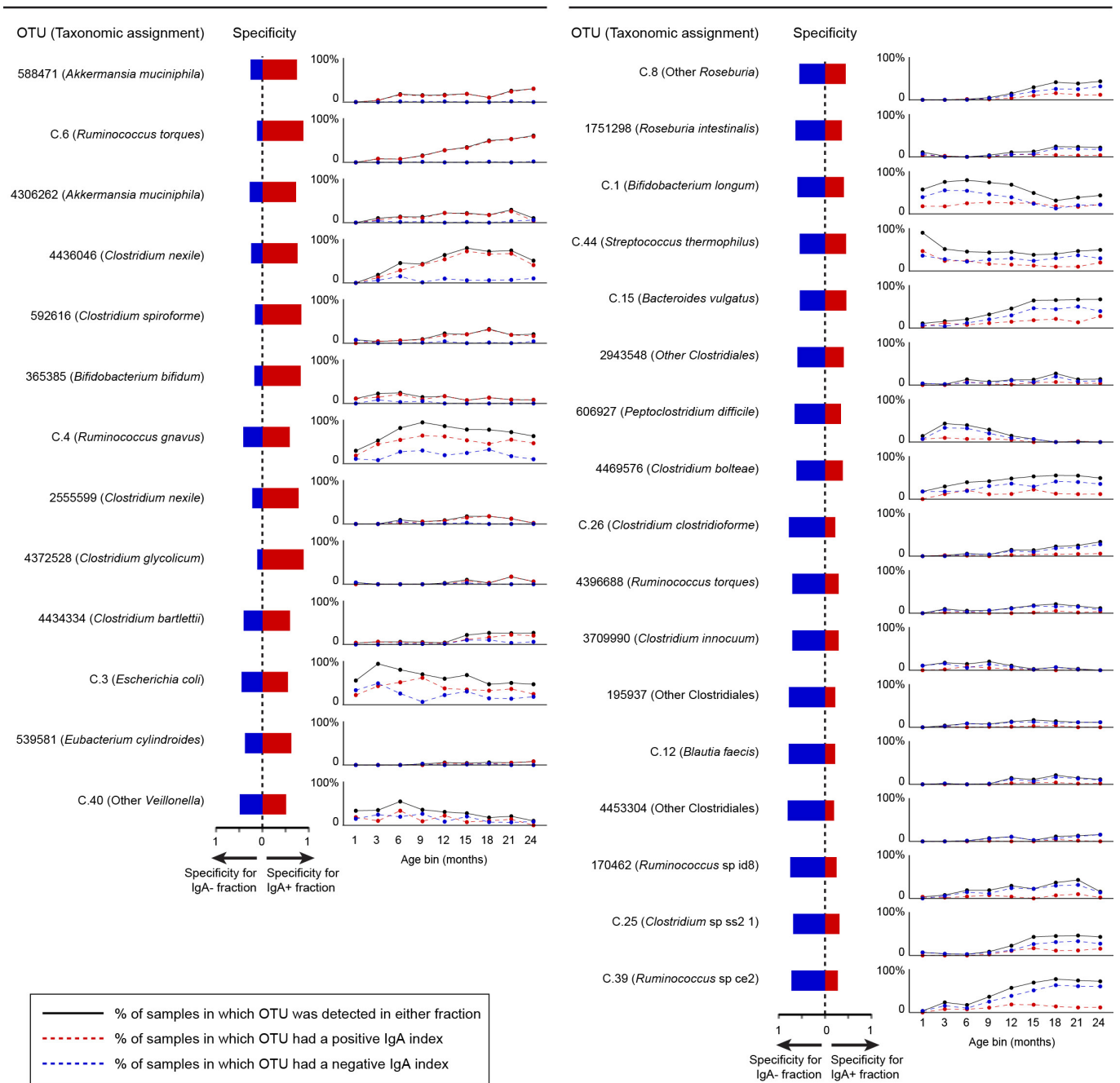


**Extended Data Figure 5 | Further characterization of IgA responses to members of the microbiota in the USA twin cohort.** **a.** Evidence that IgA indices are independent of relative abundance. IgA indices for all OTUs are plotted against their relative abundance in the 'input fraction' for all infant/child, maternal, and gnotobiotic mouse faecal samples analysed by BugFACS ( $n = 22,713$  comparisons). **b.** Mean IgA indices  $\pm$  s.e.m. for two OTUs whose IgA targeting varied significantly with age across all 80 individuals in the twin birth cohort (FDR-corrected Kruskal–Wallis test,  $P < 0.05$ ). **c.** Variance of IgA indices as a function of age. A total of 26 OTUs were detected in at least two individuals in all of the age bins surveyed. The variance in their IgA indices was then calculated and the non-parametric repeated-measures Friedman test was used to test for statistical significance ( $P < 0.0001$ ). Mean values  $\pm$  s.e.m. are plotted.



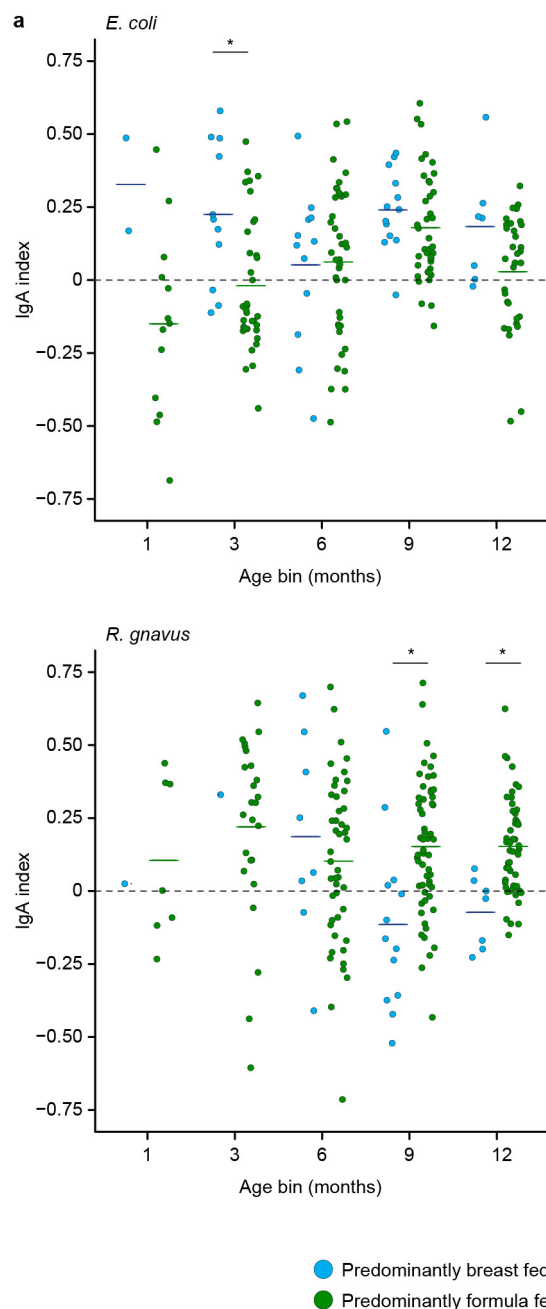
## IgA-targeted OTUs

## IgA-nontargeted OTUs



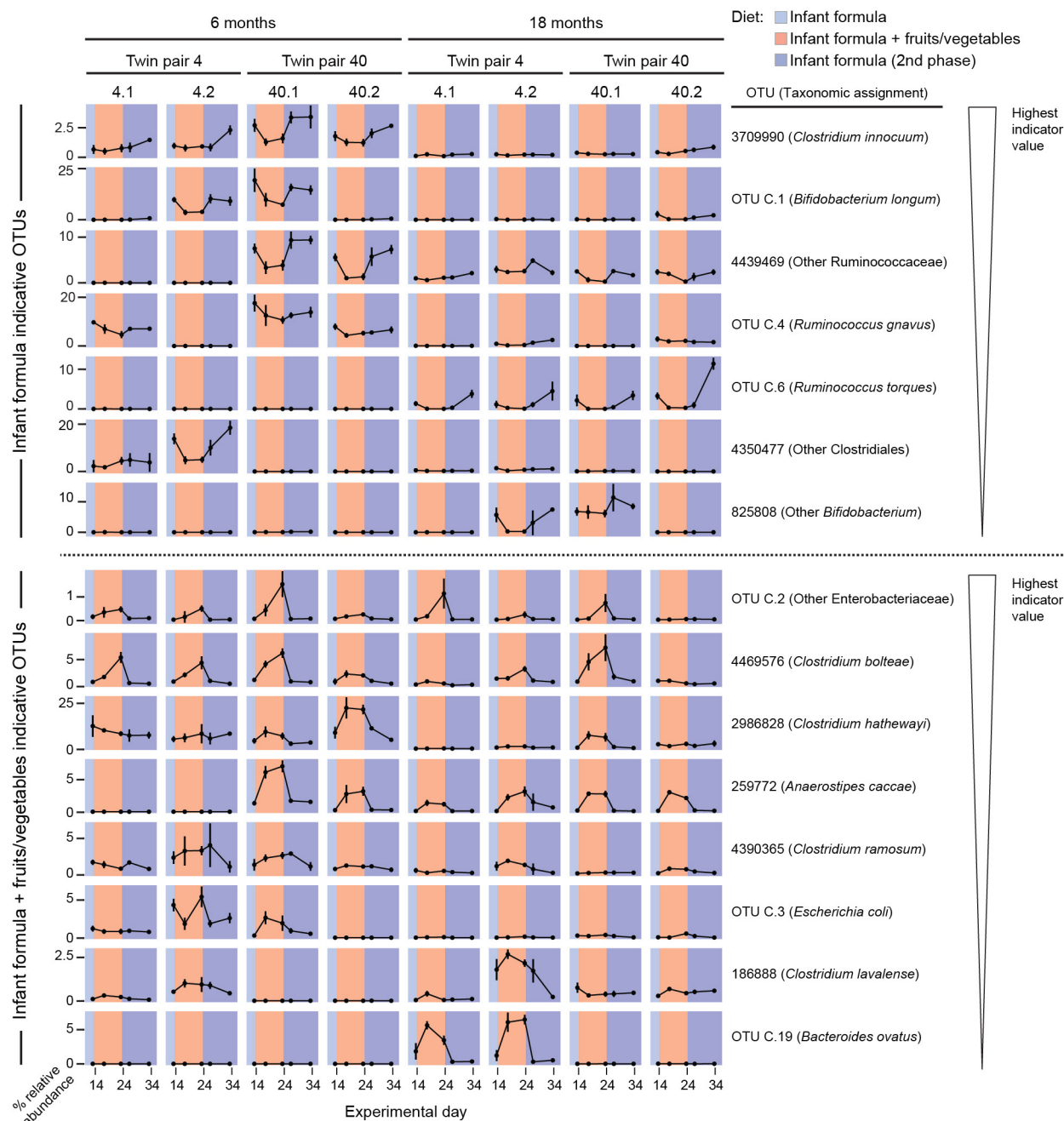
**Extended Data Figure 6 | Specificity of targeting and temporal variation in the prevalence of IgA-targeted or non-targeted taxa.** Specificity values from the indicator species analysis were calculated across all time points for the 30 OTUs identified as consistently IgA-targeted or non-targeted in Fig. 1a. Prevalence of the taxa, defined as detection in either the IgA<sup>+</sup> or IgA<sup>-</sup> fraction, was plotted against the percentage of samples in which a given taxon had a positive or negative IgA index ( $n = 4,186$  IgA index values analysed). The results reveal a group of OTUs that increased in prevalence over the course of the first 2 years of postnatal life and had very high 'specificity' for either the IgA<sup>+</sup> or IgA<sup>-</sup> fraction (that is, across the population of faecal samples, most 16S rRNA reads for a given OTU were detected in one of the two fractions). This group included *R. torques*

OTU C.6 and *C. nexile* OTU 4436046 that were IgA targeted in the majority of twins (when they are detectable in their microbiota), as well as *Ruminococcus* sp. ce2 OTU C.39 which was IgA<sup>-</sup> in most of the children. A second group of OTUs became more prevalent with age but members had a much weaker, albeit statistically significant, association with one or the other sorted fraction (for example, *R. gnavus* OTU C.4 and *B. vulgatus* OTU C.15). A third group of OTUs were highly specific for a given sorted fraction but were only detected in a minority ( $\leq 20\%$ ) of children. This last group included two strongly IgA-targeted OTUs assigned to *A. muciniphila* (OTU 588471 and OTU 4306262). Intriguingly, these two OTUs co-occurred just once among the 176 BugFACS samples in which *A. muciniphila* was detected ( $P < 0.0001$ ,  $\chi^2$  test).



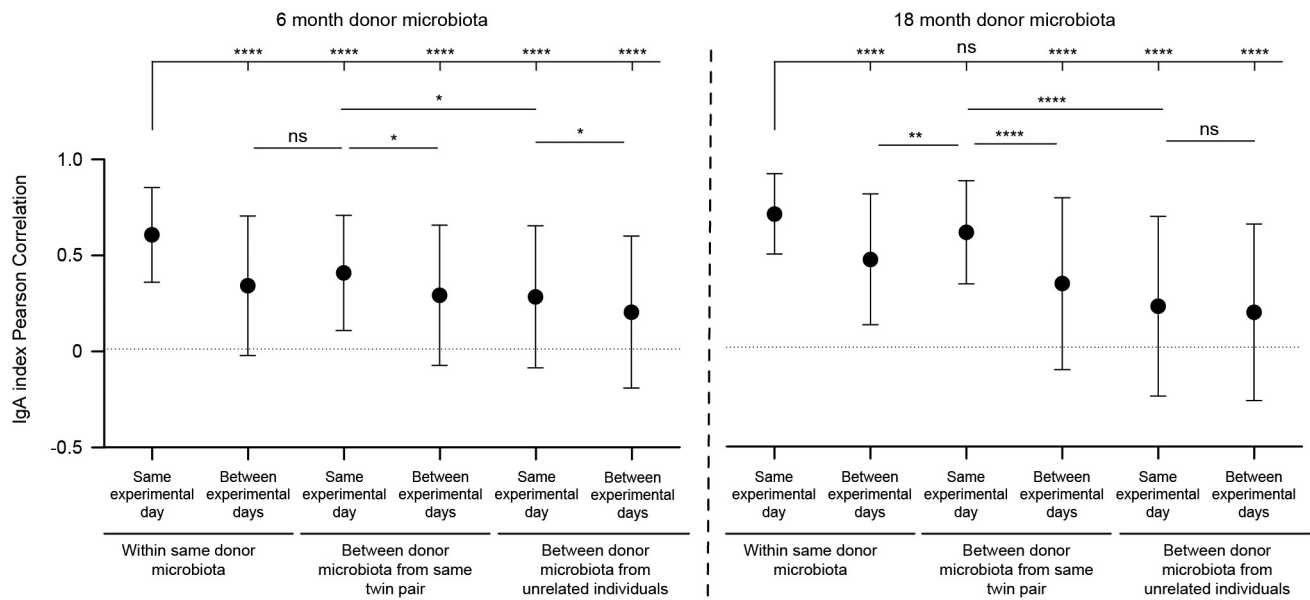
#### Extended Data Figure 7 | Effects of diet on gut mucosal IgA responses to members of the microbiota.

The analysis was constrained to those faecal samples where a diet history had been collected within 10 days of procuring the specimens ( $n = 276$ ). After FDR correction with the Benjamini–Hochberg procedure, IgA targeting of 2 of the 30 taxa identified in Fig. 1a varied significantly as a function of breast versus formula feeding. Each circle represents results from a given faecal sample. Samples are colour-coded on the basis of the type of milk diet being consumed by the donor at the time of faecal sampling. Horizontal lines in each column represent mean values.  $*P < 0.05$  (Mann–Whitney  $U$ -test of the differences between breast and formula fed).



**Extended Data Figure 8 | Diet-dependent changes in composition of the faecal microbiota of gnotobiotic mice.** Indicator species analysis was used to identify taxa from the RF-derived model of gut microbiota maturation whose abundances varied consistently by diet treatments ( $n = 9,999$  permutations with 'mouse' as the grouping variable). The top 60 ranked OTUs in the model (on the basis of their feature importance scores) were included in the analysis; those OTUs with statistically

significant diet-dependent partitioning ( $P < 0.05$ ) after FDR correction and with an indicator value  $> 0.5$  are shown, ranked from highest to lowest indicator value for infant formula-discriminatory (upper portion of figure) and 'infant formula plus fruits and vegetables'-discriminatory (lower portion of figure) (see Supplementary Table 19 for results of the indicator species analysis). Mean values for relative abundances in the faecal microbiota at each time point are plotted  $\pm$  s.d.



**Extended Data Figure 9 | Similarity in IgA responses as a function of microbiota donor, twin pair, and time after transplantation.** Pearson's correlation coefficients were calculated with IgA index data from all faecal samples collected from gnotobiotic mice that had been analysed by

BugFACS. Mean values  $\pm$  s.d. are shown for the indicated comparisons. \*\*\*\* $P < 0.0001$ ; \*\* $P < 0.01$ ; \* $P < 0.05$  (Kruskal–Wallis test with Dunn's correction for multiple comparisons;  $n = 5,029$  total comparisons).



# The Brazilian Zika virus strain causes birth defects in experimental models

Fernanda R. Cugola<sup>1\*</sup>, Isabella R. Fernandes<sup>1,2\*</sup>, Fabiele B. Russo<sup>1,3\*</sup>, Beatriz C. Freitas<sup>2</sup>, João L. M. Dias<sup>1</sup>, Katia P. Guimarães<sup>1</sup>, Cecília Benazzato<sup>1</sup>, Nathalia Almeida<sup>1</sup>, Graciela C. Pignatari<sup>1,3</sup>, Sarah Romero<sup>2</sup>, Carolina M. Polonio<sup>4</sup>, Isabela Cunha<sup>4</sup>, Carla L. Freitas<sup>4</sup>, Wesley N. Brandão<sup>4</sup>, Cristiano Rossato<sup>4</sup>, David G. Andrade<sup>4</sup>, Daniele de P. Faria<sup>5</sup>, Alexandre T. Garcez<sup>5</sup>, Carlos A. Buchpiguel<sup>5</sup>, Carla T. Braconi<sup>6</sup>, Erica Mendes<sup>6</sup>, Amadou A. Sall<sup>7</sup>, Paolo M. de A. Zanotto<sup>6</sup>, Jean Pierre S. Peron<sup>4</sup>, Alysson R. Muotri<sup>2</sup> & Patricia C. B. Beltrão-Braga<sup>1,8</sup>

Zika virus (ZIKV) is an arbovirus belonging to the genus *Flavivirus* (family *Flaviviridae*) and was first described in 1947 in Uganda following blood analyses of sentinel *Rhesus* monkeys<sup>1</sup>. Until the twentieth century, the African and Asian lineages of the virus did not cause meaningful infections in humans. However, in 2007, vectored by *Aedes aegypti* mosquitoes, ZIKV caused the first noteworthy epidemic on the Yap Island in Micronesia<sup>2</sup>. Patients experienced fever, skin rash, arthralgia and conjunctivitis<sup>2</sup>. From 2013 to 2015, the Asian lineage of the virus caused further massive outbreaks in New Caledonia and French Polynesia. In 2013, ZIKV reached Brazil, later spreading to other countries in South and Central America<sup>3</sup>. In Brazil, the virus has been linked to congenital malformations, including microcephaly and other severe neurological diseases, such as Guillain-Barré syndrome<sup>4,5</sup>. Despite clinical evidence, direct experimental proof showing that the Brazilian ZIKV (ZIKV<sup>BR</sup>) strain causes birth defects remains absent<sup>6</sup>. Here we demonstrate that ZIKV<sup>BR</sup> infects fetuses, causing intrauterine growth restriction, including signs of microcephaly, in mice. Moreover, the virus infects human cortical progenitor cells, leading to an increase in cell death. We also report that the infection of human brain organoids results in a reduction of proliferative zones and disrupted cortical layers. These results indicate that ZIKV<sup>BR</sup> crosses the placenta and causes microcephaly by targeting cortical progenitor cells, inducing cell death by apoptosis and autophagy, and impairing neurodevelopment. Our data reinforce the growing body of evidence linking the ZIKV<sup>BR</sup> outbreak to the alarming number of cases of congenital brain malformations. Our model can be used to determine the efficiency of therapeutic approaches to counteracting the harmful impact of ZIKV<sup>BR</sup> in human neurodevelopment.

The recent increase in microcephaly cases in Brazil has been associated with the outbreak of Zika virus (ZIKV)<sup>7</sup>, originating from an Asian-lineage strain that can be spread by *Aedes aegypti* mosquitoes<sup>8</sup>. The Brazilian ZIKV (ZIKV<sup>BR</sup>) has been detected in the placenta and amniotic fluid of two women with microcephalic fetuses<sup>9–11</sup> and in the blood of microcephalic newborns<sup>10,12</sup>, suggesting that the virus can cross the placental membrane. The virus has also been identified in the brains and retinas of microcephalic fetuses<sup>11–13</sup>. However, there is no direct evidence of the mechanism by which ZIKV<sup>BR</sup> causes brain malformations. A previous study revealed that the African ZIKV (ZIKV<sup>AF</sup>, strain MR-766) has the ability to infect human skin cells<sup>14</sup>. Neurons and astrocytes in the mouse brain could also be infected, inducing hippocampal degeneration and necrosis of pyriform cells

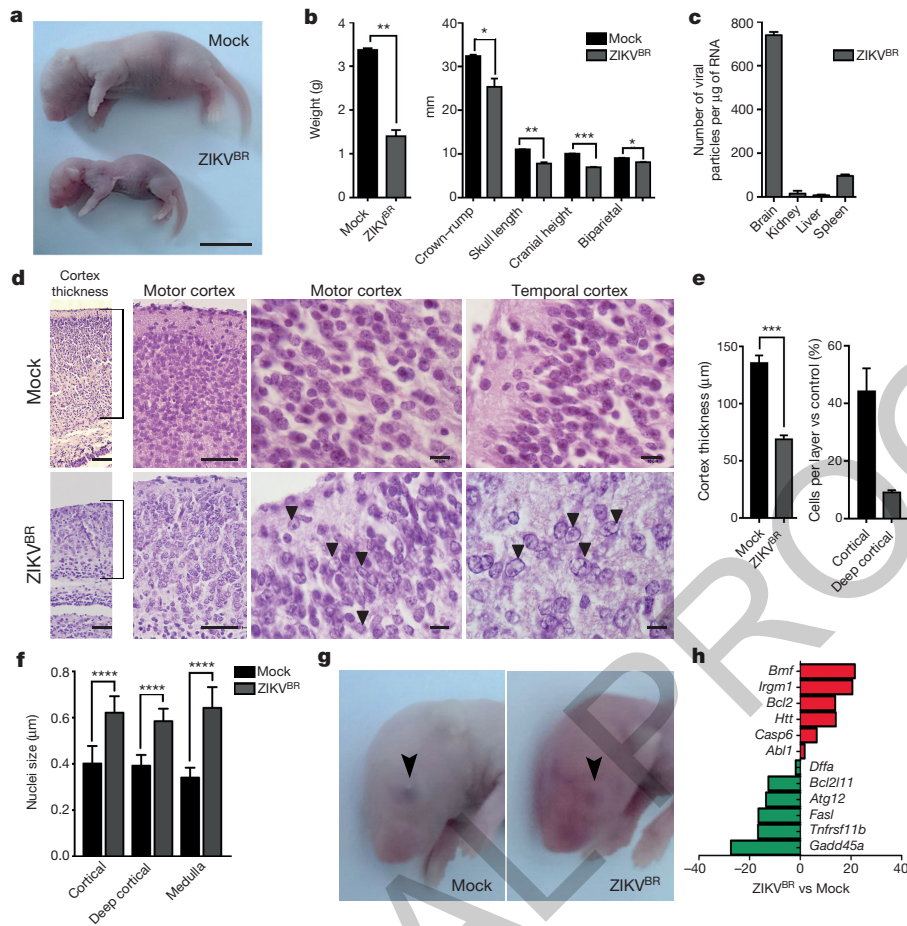
7 days post-infection<sup>15</sup>. More recently, ZIKV<sup>AF</sup> was also shown to infect human pluripotent stem cell (hPSC)-derived neural progenitor cells (NPCs) *in vitro*, which induced apoptotic cell death<sup>16</sup>. These studies were performed using the MR-766 ZIKV<sup>AF</sup> strain isolated in Uganda in 1947, which shares 87–90% sequence similarity with the Polynesian and Brazilian isolates<sup>3,17</sup>. Nevertheless, because severe congenital malformations were not reported for African isolates, there is a need to study the association of ZIKV with microcephaly and birth defects with isolates from affected localities, such as the ZIKV<sup>BR</sup> strain. Therefore, there is an urgent need to develop model systems to determine the relationship between infection with the ZIKV<sup>BR</sup> strain and birth defects.

We used ZIKV<sup>BR</sup> isolated from a febrile case in the state of Paraíba, in the northeast of Brazil in 2015 (see Methods). To evaluate the causal relationship between ZIKV<sup>BR</sup> and birth defects, including brain malformation during development, we first used a murine experimental model in which SJL and C57BL/6 pregnant mice were infected with ZIKV<sup>BR</sup>, evaluating newborns immediately after birth (Extended Data Fig. 1a). Notably, similar to ZIKV<sup>BR</sup>-infected human newborns<sup>18,19</sup>, pups born from the SJL ZIKV<sup>BR</sup>-infected pregnant females displayed clear evidence of whole-body growth delay or intra-uterine growth restriction (IUGR)<sup>20</sup> compared to pups born from the mock-infected controls (Fig. 1a, b). Using a qPCR assay, we confirmed the presence of ZIKV<sup>BR</sup> genomic RNA in several tissues of newborn animals, observing significantly more viral RNA in the brain, confirming the neurotropic nature of the virus (Fig. 1c).

Microcephaly is perhaps the most dramatic of the birth defects reported in ZIKV<sup>BR</sup>-infected newborns<sup>4,18,19</sup>. Mouse models often fail to reproduce the severely reduced brain size and pathological alterations found in human patients<sup>21,22</sup>, probably owing to significant differences in gestation time and brain development between the two species. Nevertheless, upon close inspection of the ZIKV<sup>BR</sup>-infected mice brains, we noticed cortical malformations in the surviving animals, with reduced cell number and cortical layer thickness, signs associated with microcephaly in humans (Fig. 1d–f). At a cellular level, the neurons in the cortex, thalamus and hypothalamus displayed a ‘vacuolar nuclei’ appearance. This morphology was characterized by central emptiness and marginalized chromatin pattern with nuclear debris, suggesting ongoing cellular death (Fig. 1d and Extended Data Fig. 2). In addition, we also noticed apparent ocular abnormalities, reminiscent of that observed in human patients<sup>23</sup> (Fig. 1g). Thus, SJL infected pups presented congenital malformations compatible with ZIKV<sup>BR</sup>-infected human newborns. While the impact of ZIKV<sup>BR</sup>

<sup>1</sup>University of São Paulo, Department of Surgery, Stem Cell Laboratory, São Paulo, São Paulo 05508-270, Brazil. <sup>2</sup>University of California San Diego, School of Medicine, Department of Pediatrics/Rady Children's Hospital San Diego, Department of Cellular & Molecular Medicine, Stem Cell Program, La Jolla, California 92037-0695, USA. <sup>3</sup>Tismoo, The Biotech Company, São Paulo, São Paulo 01401-000, Brazil. <sup>4</sup>University of São Paulo, Department of Immunology, Neuroimmune Interactions Laboratory, São Paulo, São Paulo 05508-000, Brazil. <sup>5</sup>University of São Paulo, Department of Radiology and Oncology, USP School of Medicine, São Paulo, São Paulo 05403-010, Brazil. <sup>6</sup>University of São Paulo, Department of Microbiology, Institute of Microbiology Sciences, Laboratory of Molecular Evolution and Bioinformatics, São Paulo, São Paulo 05508-000, Brazil. <sup>7</sup>Institute Pasteur in Dakar, Dakar 220, Sénégal. <sup>8</sup>University of São Paulo, School of Arts, Sciences and Humanities, Department of Obstetrics, São Paulo, São Paulo 03828-000, Brazil.

\*These authors contributed equally to this work.



**Figure 1 | ZIKV<sup>BR</sup> infection in SJL mice.** **a**, SJL pups born with IUGR. Scale bar, 1 cm. **b**, Total body weight, crown-rump and skull measurements in pups born from infected animals ( $n=6$  pups, comprising 3 mice from 2 separate litters; error bars, s.e.m.;  $t$ -test,  $^{**}P < 0.01$ ). **c**, ZIKV<sup>BR</sup> RNA detected in SJL pup tissues ( $n=6$  pups, comprising 3 mice from 2 separate litters; error bars, s.e.m.;  $t$ -test). **d**, Histopathological aspect of the cortical organization (brackets) in infected brains, including intranuclear vacuoles, and 'empty' nuclei aspect with chromatin margination in neurons (arrowheads). Scale bar, 100  $\mu$ m (left panels), 50  $\mu$ m (middle panels) and 10  $\mu$ m (right panels). **e**, ZIKV<sup>BR</sup>-infected brains

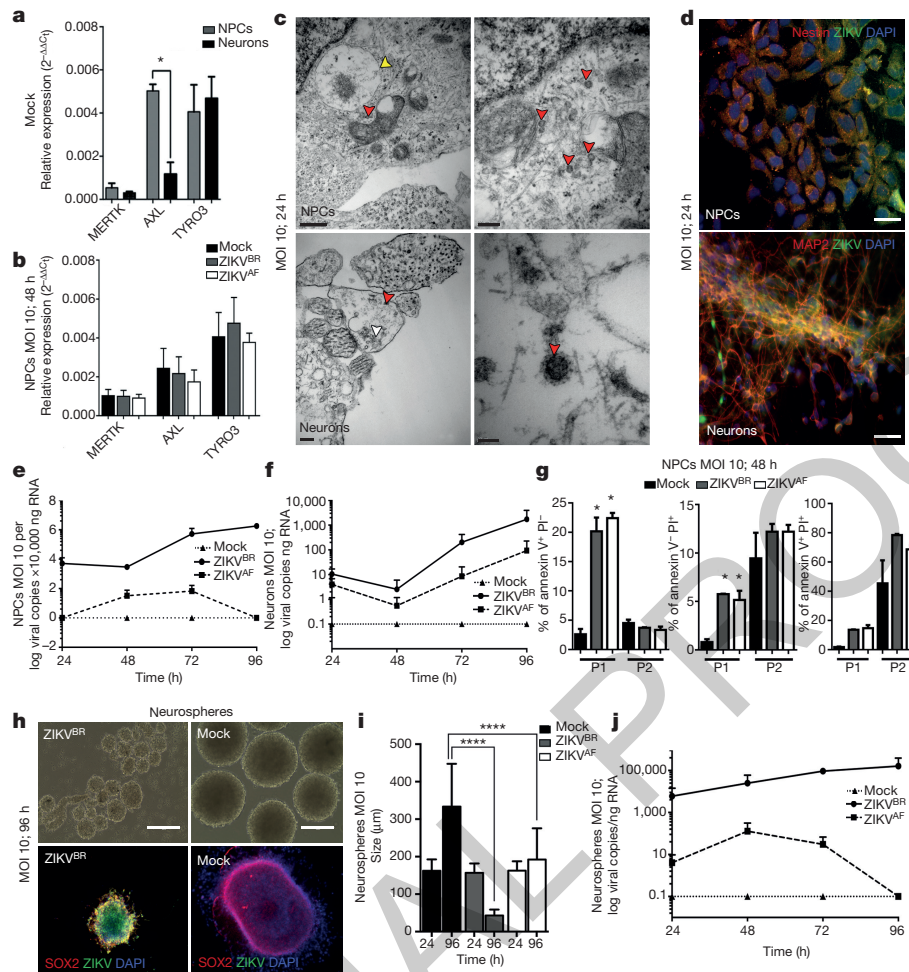
displayed a reduced cortical layer thickness ( $n=6$  pups, comprising 3 mice from 2 separate litters; error bars, s.e.m.;  $t$ -test,  $^{***}P < 0.001$ ). Infected brains have fewer cells per layer ( $n=6$  pups, comprising 3 mice from 2 separate litters; error bars, s.e.m.;  $t$ -test,  $^{**}P < 0.1$ ). **f**, ZIKV<sup>BR</sup>-infected cortical neurons have pronounced nuclei (diameter) (cortical,  $n=31$ ; deep cortical,  $n=21$  and medulla,  $n=41$  nuclei; error bars, s.e.m.; two-way ANOVA,  $^{****}P < 0.001$ ). **g**, Ocular malformations (arrow) in the ZIKV<sup>BR</sup>-infected pups. **h**, Cell death gene expression signature in the brains of ZIKV<sup>BR</sup>-infected pups ( $n=2$  mice per group; threshold = twofold).

infection in the SJL mice was notable, no major body alterations were detected in pups from the infected C57BL/6 animals (Extended Data Fig. 1a–c). To exclude potential minor alterations in the C57BL/6 mice, computed tomography (CT) scans were performed to quantify the skull/body volumes. No significant changes in the pups born from the ZIKV<sup>BR</sup>-infected C57BL/6 mothers were observed compared to the controls (Extended Data Fig. 1d, e). A diagnostic qPCR assay of the pups from the ZIKV<sup>BR</sup>-infected animals was negative, suggesting that the virus did not cross the placenta in the C57BL/6 mouse strain (Extended Data Fig. 1f). To elucidate the type of cell death induced by ZIKV<sup>BR</sup> in the brains of SJL pups, we used a qPCR array to distinguish different molecular pathways involved. Our data clearly indicate that ZIKV<sup>BR</sup> infection influence the regulation of genes intimately linked to autophagy and apoptosis, such as upregulation of *Bmf*, *Irgm1*, *Bcl2*, *Htt*, *Casp6* and *Abl1*. Conversely, *Gadd45a*, *Tnfrsf11b*, *Fas*, *Atg12*, *Bcl2l1* and *Dffa* were highly suppressed (Fig. 1h and Extended Data Fig. 1g, h).

Next, we evaluated the impact of ZIKV<sup>BR</sup> infection in human neural cells derived from hPSCs to establish a correlation between ZIKV and impairment of neurogenesis (Extended Data Fig. 3a). We generated human cortical NPCs and neurons from healthy donor hPSCs. First, we determined the expression levels of the TYRO3, AXL and MERTK (TAM) receptors tyrosine kinases in NPCs and neurons. This is an

important family of receptors used for cell invasion by the Dengue virus and ZIKV, and AXL has been recently proposed as a candidate receptor for ZIKV infection during neurogenesis<sup>14,24</sup>. Mock-infected NPCs expressed higher levels of AXL when compared to mock-infected neurons (Fig. 2a). However, no significant changes in expression levels were observed upon ZIKV infection in NPCs (Fig. 2b). We then investigated the impact of ZIKV<sup>BR</sup> and ZIKV<sup>AF</sup> infection in NPCs and neurons. After infection using a viral multiplicity of infection (MOI) of 10, ZIKV<sup>BR</sup> particles were detected inside the NPCs and neurons at several stages of viral assembly using transmission electron microscopy (Fig. 2c). Immunostaining performed on NPCs and neurons at both an MOI of 10 and an MOI of 1 revealed production of viral protein aggregates (Fig. 2d and Extended Data Fig. 3b, c). With an MOI of 10, the amount of ZIKV<sup>BR</sup> particles in the NPC and neuron culture supernatant increased over time, suggesting the efficient production of infectious viral particles (Fig. 2e, f). With an MOI of 1, NPCs, but not neurons, continued to produce ZIKV<sup>BR</sup> RNA in the culture supernatant (Extended Data Fig. 3d, e). After 96 h post-infection we observed significant cell death in NPC cultures using fluorescence-activated cell (FAC) analyses. We quantified cell death over time in NPC cultures and detected an increase in the number of apoptotic/necrotic cells both in the ZIKV<sup>BR</sup>- and ZIKV<sup>AF</sup>-treated cultures compared to the





**Figure 2 | ZIKV infection *in vitro*.** **a**, Relative expression of TAM receptors ( $n = 2$  technical replicates from two pooled different donors; error bars, s.e.m.;  $t$ -test; \*\*\* $P < 0.01$ ). **b**, Expression of TAM receptors in NPCs after ZIKV<sup>BR</sup> infection (MOI = 10) at 48 h post-infection ( $n = 2$  technical replicates from two pooled different donors; error bars, s.e.m.; one-way ANOVA,  $P > 0.05$ ). **c**, TEM detection of ZIKV<sup>BR</sup> viral particles 24 h post-infection at MOI = 10 (red arrowheads) inside NPCs (top panels) and neurons (bottom panels). Yellow arrowheads, viral factories; white arrowheads, immature viral particles. Scale bars/magnifications, 0.5  $\mu\text{m}$ /40,000 $\times$  (top left); 200 nm/80,000 $\times$  (top right); 0.2  $\mu\text{m}$ /50,000 $\times$  (bottom left); 50 nm/300,000 $\times$  (bottom right). **d**, Immunofluorescence revealed susceptibility to infection in NPC and neurons with the ZIKV<sup>BR</sup> (MOI = 10) at 24 h post-infection. Scale bar, 25  $\mu\text{m}$ . **e**, ZIKV<sup>BR</sup> replication dynamics in NPCs (MOI = 10;  $n = 2$

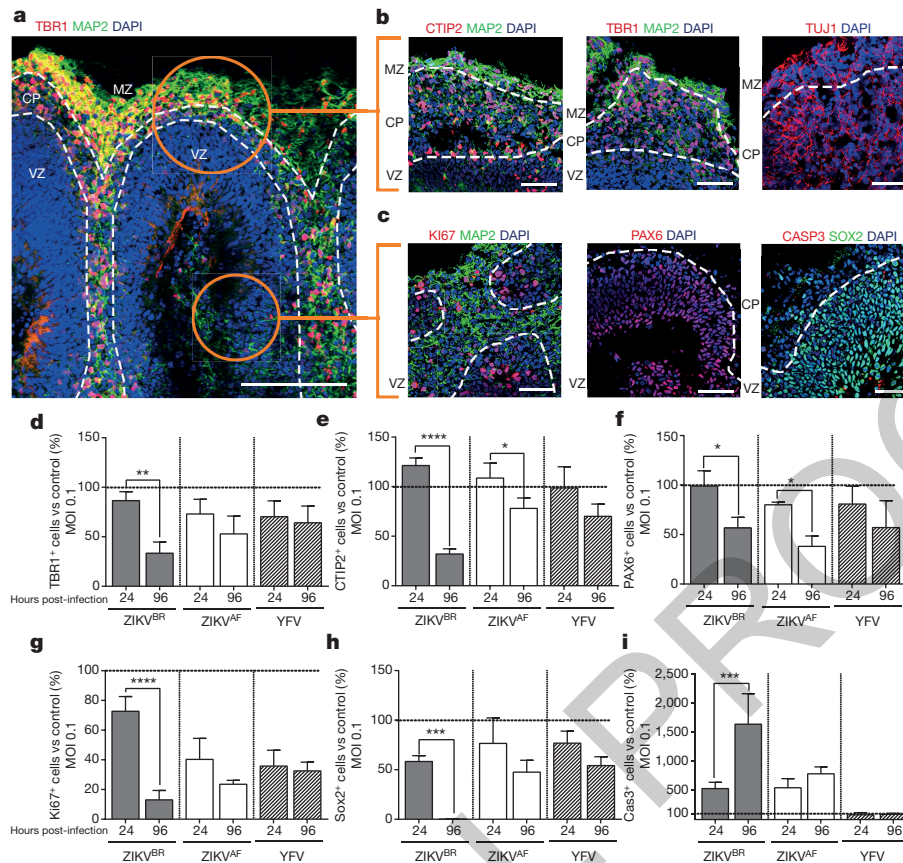
technical replicates from RNA of two different donors). **f**, ZIKV<sup>BR</sup> replication dynamics in neurons (MOI = 10;  $n = 2$  technical replicates from RNA of two different donors). **g**, NPC death measured by FACS with two different cell gating sizes (P1 and P2). Apoptosis (left panel), necrosis (middle panel) and late apoptosis (right panel) (MOI = 10; 48 h post-infection;  $n = 2$  technical replicates from two different donors; error bars, s.e.m.; two-way ANOVA, \* $P < 0.05$ ). P1, propidium iodide. **h**, Representative images of human neurospheres infected with ZIKV<sup>BR</sup> (MOI = 10; 96 h post-infection). Scale bar, 200  $\mu\text{m}$ . **i**, Alterations in neurosphere diameter over time (MOI = 10;  $n = 25$  neurospheres from two different donors for each time point; error bars, s.e.m.; ANOVA, \*\*\*\* $P < 0.0001$ ). **j**, ZIKV replication dynamics in neurospheres (MOI = 10;  $n = 2$  technical replicates from two different donors).

mock-infected cultures at an MOI of 10 (Fig. 2g), but not at an MOI of 1 during the same time frame (Extended Data Fig. 3f). No difference in neuronal cell death was observed between the two ZIKV strains at an MOI of 10 and an MOI of 1 (Extended Data Fig. 3g, h).

Next, we challenged two three-dimensional neural cell culture systems, neurospheres and cerebral organoids, with ZIKV<sup>BR</sup> and ZIKV<sup>AF</sup>. We generated neurospheres by growing human NPCs in suspension. While the mock-infected control neurospheres continued to grow over time, the ZIKV<sup>BR</sup>-infected neurospheres (MOI of 10) displayed evident morphological abnormalities with signs of cell death (Fig. 2h). The sizes of the neurospheres infected with ZIKV<sup>BR</sup> were significantly smaller than the mock-control and ZIKV<sup>AF</sup>-infected at 96 h post-infection (Fig. 2i). A less dramatic effect is observed at an MOI of 1, where both ZIKV<sup>BR</sup> and ZIKV<sup>AF</sup> infection reduced the size of the neurospheres compared to mock-infected controls (Extended Data Fig. 4a, b). These observations were paired with increased ZIKV<sup>BR</sup> replication in these cultures at both MOIs of 10 and 1 (Fig. 2j and Extended Data Fig. 4c).

These results suggest that ZIKV<sup>BR</sup> induces cell death in human NPCs, impairing the growth and morphogenesis of healthy neurospheres (Extended Data Fig. 4d–f).

The majority of the described cases of ZIKV<sup>BR</sup>-infected newborns (95%) had malformations of cortical development<sup>4</sup>. Thus, we also used brain organoids generated from hPSCs and human embryonic stem cells, to evaluate the impact of ZIKV<sup>BR</sup> on human cortical development. In the following experiments, alongside the ZIKV<sup>AF</sup> and mock infection, we added the Yellow Fever virus (YFV), a slow replicating attenuated live-vaccine Flavivirus that has a low risk of causing neuropathy. Cerebral organoids are three-dimensional, self-organized, stem-cell-derived models that recapitulate the first trimester of human neurodevelopment, including the molecular and cellular architecture reminiscent of the fetal cortex<sup>25</sup>. Organoids show some degree of lamination and resembled the human neocortex in terms of the spatial relationships of the progenitor populations, defined here as a proliferative ventricular



**Figure 3 | Cortical alterations in human brain organoids infected with ZIKV.** **a**, Representative image of a human cerebral organoid showing the marginal zone (MZ), cortical plate (CP) and ventricular zone (VZ), delineated by dotted white lines. Scale bar, 200  $\mu$ m. **b**, Representative images of the CP stained for CTIP2, TBR1, MAP2 or TUJ1 (neurons). Scale bar, 50  $\mu$ m. **c**, Representative images of the proliferative regions in the VZ stained for KI67, PAX6, and cleaved caspase-3 (CASP3). Scale bar, 50  $\mu$ m. **d**, Percentage of TBR1-positive cells in relation to mock-infected controls (dotted line) (MOI = 0.1;  $n = 3$  replicates from three human cell lines; error bars, s.e.m.; ANOVA, \*\* $P = 0.0025$ ). **e**, Percentage

zone, post-migratory neurons in cortical plate and a marginal zone (Fig. 3a–c). We infected organoids with ZIKV<sup>BR</sup>, ZIKV<sup>AF</sup> and YFV using an MOI of 0.1 and compared to mock-infected organoids at 24 and 96 h post-infection. We quantified the percentage of different subtypes of cortical neurons, TBR1-positive or CTIP2-positive cells (deep-layer V/V1) in the cortical plate, finding a significant reduction in their number and respective cortical plate thickness in the ZIKV<sup>BR</sup>-infected organoids compared to the others. A significant reduction in TBR1-positive cells was observed in the ZIKV<sup>BR</sup>-infected organoids at 96 h post-infection, while CTIP2-positive cells were significantly reduced in both ZIKV<sup>BR</sup>- and ZIKV<sup>AF</sup>-infected organoids at the same time point (Fig. 3d, e and Extended Data Fig. 5a–f). Consistent with the reduced population of cortical neurons, we observed a significant decrease in PAX6-positive cells (dorsal forebrain progenitor cells) following ZIKV infection (Fig. 3f and Extended Data Fig. 5d). Dividing cells in the ventricular zone, detected by the population of KI67- and SOX2-positive cells, were only significantly reduced in the ZIKV<sup>BR</sup>-infected organoids (Fig. 3g, h and Extended Data Fig. 5d). As observed in our other *in vitro* models, the number of apoptotic cells (cleaved caspase 3- and TUNEL-positive cells) was increased in organoids infected with ZIKV<sup>BR</sup>, possibly explaining the decrease in the NPC population (Fig. 3i and Extended Data Fig. 5g, h).

ZIKV<sup>AF</sup> was derived from a zoonotic agent associated with primates in Africa, whereas ZIKV<sup>BR</sup> is an isolate from a lineage adapted to human-to-human transmission for the past 70 years. As an entry

point to establishing the potential mechanistic adaptive differences between ZIKV<sup>BR</sup> and ZIKV<sup>AF</sup> towards human cells, we also generated brain organoids from non-human primate (chimpanzee) pluripotent stem cells. We infected these chimpanzee's cerebral organoids (MOI of 0.1) and measured the impact on cortical neurons at 24 and 96 h post-infection. ZIKV<sup>BR</sup> failed to reduce the percentage of either TBR1- or CTIP2-positive cells in non-human primates (Extended Data Fig. 5i, j). Consistently, the kinetics of infection were different between the two ZIKV isolates. While ZIKV<sup>BR</sup> did not replicate in the chimpanzee organoids, ZIKV<sup>AF</sup> seemed well adapted to these cells (Extended Data Fig. 5k).

To evaluate the causal relationship between ZIKV congenital infection and birth defects, we used a murine experimental model, in which pregnant SJL and C57BL/6 mice were infected with ZIKV<sup>BR</sup>. Notably, the SJL strain was susceptible to viral infection of fetal tissues, causing severe IUGR that resembled the affected Brazilian newborns, including signs of microcephaly, such as cortical malformations<sup>4</sup>. We also showed that ZIKV<sup>BR</sup> induced apoptosis and autophagy in the mouse neural tissue. This is in accordance with the literature, as it has been previously demonstrated that ZIKV induces and localizes in autophagic phagosomes<sup>14</sup>. To our knowledge, this is the first report showing a gene expression profile that correlates to cell death in the brains of microcephalic newborn ZIKV<sup>BR</sup>-infected mice, corroborating a causal relationship. It is unclear why the virus could not cross the placenta of C57BL/6 mice, but this result may be due to the robust anti-viral



immune response of this mouse strain, which secretes significant levels of type I/II interferon, known to confer resistance to ZIKV<sup>14,26</sup>. These data suggest that genetic differences could explain in part why some ZIKV-infected pregnant women give birth to newborns without detectable congenital brain malformations<sup>27</sup>. Nonetheless, our murine model is a valuable tool for future pre-clinical studies, such as vaccine development. The presence of major cortical histological abnormalities in the pups with IUGR prompted us to use an hPSC model to study the impact of ZIKV on neurodevelopment. ZIKV infects cells at different stages of brain maturation leading to alterations in the cortical layer organization. While this manuscript was under review, two other papers revealed the impact of previously established ZIKV strains on human organoids, confirming our observations with ZIKV<sup>BR</sup> (refs 28,29). Finally, our data using a non-human primate organoids suggested that the ZIKV<sup>BR</sup> might have experienced adaptive changes in human cells. In fact, it has been demonstrated that the Asian lineage of ZIKV is undergoing codon usage adaptation towards biases observed in highly expressed human genes<sup>30</sup>. Our findings support the hypothesis that microcephaly is a distinctive feature of recent ZIKV Asian-lineage virus, which originated in the Pacific and is now spreading in South and Central America.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 8 March 2015; accepted 4 May 2016.**

**Published online 11 May 2016.**

- Dick, G. W., Kitchen, S. F. & Haddow, A. J. Zika virus. I. Isolations and serological specificity. *Trans. R. Soc. Trop. Med. Hyg.* **46**, 509–520 (1952).
- Lancioti, R. S. et al. Genetic and serologic properties of Zika virus associated with an epidemic, Yap State, Micronesia, 2007. *Emerg. Infect. Dis.* **14**, 1232–1239 (2008).
- Faria, N. R. et al. Zika virus in the Americas: early epidemiological and genetic findings. *Science* **352**, 345–349 (2016).
- de Fatima Vasco Aragao, M. et al. Clinical features and neuroimaging (CT and MRI) findings in presumed Zika virus related congenital infection and microcephaly: retrospective case series study. *Br. Med. J.* **353**, i1901 (2016).
- Cao-Lormeau, V. M. et al. Guillain-Barré Syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study. *Lancet* **387**, 1531–1539 (2016).
- Rasmussen, S. A., Jamieson, D. J., Honein, M. A. & Petersen, L. R. Zika virus and birth defects - reviewing the evidence for causality. *N. Engl. J. Med.* <http://dx.doi.org/10.1056/NEJMs1604338> (2016).
- Campos, G. S., Bandeira, A. C. & Sardi, S. I. Zika virus outbreak, Bahia, Brazil. *Emerg. Infect. Dis.* **21**, 1885–1886 (2015).
- Hayes, E. B. Zika virus outside Africa. *Emerg. Infect. Dis.* **15**, 1347–1350 (2009).
- Sarno, M. et al. Zika Virus infection and stillbirths: a case of hydrops fetalis, hydranencephaly and fetal demise. *PLoS Negl. Trop. Dis.* **10**, e0004517 (2016).
- Calvet, G. et al. Detection and sequencing of Zika virus from amniotic fluid of fetuses with microcephaly in Brazil: a case study. *Lancet Infect. Dis.* [http://dx.doi.org/10.1016/S1473-3099\(16\)00095-5](http://dx.doi.org/10.1016/S1473-3099(16)00095-5) (2016).
- Martines, R. B. et al. Notes from the field: evidence of Zika virus infection in brain and placental tissues from two congenitally infected newborns and two fetal losses - Brazil, 2015. *MMWR Morb. Mortal. Wkly. Rep.* **65**, 159–160 (2016).
- Malakar, J. et al. Zika virus associated with microcephaly. *N. Engl. J. Med.* **374**, 951–958 (2016).
- Ventura, C. V., Maia, M., Bravo-Filho, V., Góis, A. L. & Belfort, R., Jr. Zika virus in Brazil and macular atrophy in a child with microcephaly. *Lancet* **387**, 228 (2016).
- Hamel, R. et al. Biology of Zika virus infection in human skin cells. *J. Virol.* **89**, 8880–8896 (2015).
- Bell, T. M., Field, E. J. & Narang, H. K. Zika virus infection of the central nervous system of mice. *Arch. Gesamte Virusforsch.* **35**, 183–193 (1971).
- Tang, H. et al. Zika virus infects human cortical neural progenitors and attenuates their growth. *Cell Stem Cell* <http://dx.doi.org/10.1016/j.stem.2016.02.016> (2016).
- Faye, O. et al. Molecular evolution of Zika virus during its emergence in the 20(th) century. *PLoS Negl. Trop. Dis.* **8**, e2636 (2014).

- Brasil, P. et al. Zika virus outbreak in Rio de Janeiro, Brazil: clinical characterization, epidemiological and virological aspects. *PLoS Negl. Trop. Dis.* **10**, e0004636 (2016).
- Brasil, P. et al. Zika virus infection in pregnant women in Rio de Janeiro - preliminary report. *N. Engl. J. Med.* <http://dx.doi.org/10.1056/NEJMoa1602412> (2016).
- Jang, E. A., Longo, L. D. & Goyal, R. Antenatal maternal hypoxia: criterion for fetal growth restriction in rodents. *Front. Physiol.* **6**, 176 (2015).
- Lizarraga, S. B. et al. Cdk5rap2 regulates centrosome function and chromosome segregation in neuronal progenitors. *Development* **137**, 1907–1917 (2010).
- Pulvers, J. N. et al. Mutations in mouse *Aspm* (abnormal spindle-like microcephaly associated) cause not only microcephaly but also major defects in the germline. *Proc. Natl Acad. Sci. USA* **107**, 16595–16600 (2010).
- de Paula Freitas, B. et al. Ocular findings in infants with microcephaly associated with presumed Zika virus congenital infection in Salvador, Brazil. *JAMA Ophthalmol.* <http://dx.doi.org/10.1001/jamaophthol.2016.0267> (2016).
- Nowakowski, T. J. et al. Expression analysis highlights AXL as a candidate Zika virus entry receptor in neural stem cells. *Cell Stem Cell* <http://dx.doi.org/10.1016/j.stem.2016.03.012> (2016).
- Lancaster, M. A. et al. Cerebral organoids model human brain development and microcephaly. *Nature* **501**, 373–379 (2013).
- Rossi, S. L. et al. Characterization of a novel murine model to study Zika virus. *Am. J. Trop. Med. Hyg.* <http://dx.doi.org/10.4269/ajtmh.16-0111> (2016).
- Cauchemez, S. et al. Association between Zika virus and microcephaly in French Polynesia, 2013–15: a retrospective study. *Lancet* [http://dx.doi.org/10.1016/S0140-6736\(16\)00651-6](http://dx.doi.org/10.1016/S0140-6736(16)00651-6) (2016).
- Garcez, P. P. et al. Zika virus impairs growth in human neurospheres and brain organoids. *Science* <http://dx.doi.org/10.1126/science.aaf6116> (2016).
- Qian, X. et al. Brain-region-specific organoids using mini-bioreactors for modeling ZIKV exposure. *Cell* **165**, 1–17 (2016).
- Freire, C. C. d. M. et al. Spread of the pandemic Zika virus lineage is associated with NS1 codon usage adaptation in humans. *BioRxiv* <http://biorxiv.org/content/early/2015/11/25/032839>.

**Acknowledgements** This work was supported by grants from the Zika Network FAPESP projects 2011/18703-2 and 2014/17766-9, the NGO 'The Tooth Fairy Project', the National Institutes of Health through the U19MH107367, R01MH094753 and an NARSAD Independent Investigator Grant to A.R.M. We would like to thank M. Gonçalves and M. A. Miglino for scientific discussions, G. Amarante-Mendes for the FITC-Annexin V reagent, M. Zatz for the SJL mice, E. Durigon and his group for the ZIKV<sup>BR</sup> aliquots, P. Vasconcelos for providing a lyophilized ZIKV<sup>BR</sup> seed, N. Olsen for reagent and equipment support, D. Candido for analysis support, F. Mastrantonio for drawings and R. E. Grassi for electron and confocal microscopy support.

**Author Contributions** F.R.G. derived the NPCs, neurons and neurospheres, performed the viral infections and cell analyses and analysed the data. I.R.F. performed the viral infections of cells, processed and analysed infected brain organoids, prepared manuscript figures and analysed the data. F.B.R. derived the NPCs, performed immunostainings and analyses, prepared manuscript figures and analysed the data. B.C.F. revised the manuscript and with S.R. generated the organoid cultures from iPSCs and assisted with the immunofluorescence staining and experimental design. J.L.M.D. performed macroscopic and microscopic staining and analyses of the mice. K.O.P.G. performed the TEM experiments, RNA extraction and quantification and histopathological analyses. C.B. and N.S. performed RNA extraction and quantification and prepared figures. G.C.P. performed cell cultures, analysed the data and revised the manuscript. C.M.P., I.C., C.L.F., W.N.B. and C.R. performed cell death qPCR assays and flow cytometry. D.G.A. performed flow cytometry staining protocols and analysed the data. C.M.P., I.C. and D.G.A. infected and observed the pregnant mice daily. C.M.P., C.L.F., I.C. and C.R. processed the mouse tissues for virus quantification and further analyses. D.P.F., A.T.G. and C.A.B. performed the CT scans and analysed and plotted the data. C.T.B. and E.A.M. performed virus amplification, titration and gene expression analysis. A.A.S. provided MR766 and YFV-17D isolates and serological reagents. P.M.A.Z. designed the experiments and revised the manuscript. J.P.S.P., A.R.M. and P.C.B.B.-B. designed the experiments, analysed the data and wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.C.B.B.-B. ([patriciaacbbbraga@usp.br](mailto:patriciaacbbbraga@usp.br)); A.R.M. ([muotri@ucsd.edu](mailto:muotri@ucsd.edu)) and J.P.S.P. ([jeanpierre@usp.br](mailto:jeanpierre@usp.br)).

## METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment

**Viral culture and amplification.** A lyophilized ZIKV isolate from a clinical case in Brazil (ZIKV<sup>BR</sup>), provided by the Evandro Chagas Institute in Belém, Pará, was reconstituted in 0.5 ml of sterile DEPC water. The African-lineage MR-766 (ZIKV<sup>AF</sup>), a reference strain isolated in Uganda in 1947 and the Yellow Fever vaccine strain (YFV-17D)<sup>31</sup>, both used here as controls, were provided by the Institute Pasteur in Dakar, Senegal. *Aedes albopictus* mosquito cells (C6/36 cells) were previously prepared to culture the three viruses. C6/36 cell culture was maintained using Leibovitz's L-15 medium supplemented with 10% fetal bovine serum (FBS) (Gibco), 1% non-essential amino acids (Gibco), 1% sodium pyruvate (Gibco), 1% penicillin/streptomycin (Gibco), 0.05% of amphotericin B (Gibco) and kept at 27°C in the absence of CO<sub>2</sub>. After reaching an approximately 70% confluent monolayer, 50 µl of each viral sample was inoculated into C6/36 with an hour of adsorption, with gentle shaking every 10 min to allow the homogeneous adsorption of the viruses. At the end of the adsorption period, 5 ml of the culture media was added, plus 2% FBS, 1% non-essential amino acids and 1% sodium pyruvate. The cultures were then incubated under the same adsorption conditions. In the first subculture (T1), the infected cells were less confluent compared to the control cells but had few noticeable morphological changes. On the fourth day after infection, the second subculture (T2) was made blindly by transferring 500 µl of the T1 supernatant, followed by the third subculture, which was collected on the eighth day after infection (T3). Pronounced cytopathic effects were perceived beginning at T2. The supernatants were harvested, titrated and T3 was used for the experimental inoculation.

**Virus titration.** Titration (in PFU ml<sup>-1</sup>) of each C6/36 subculture was obtained by plaque assay to determine the amount of infectious viral particles (PFU). The virus titration was performed in porcine kidney epithelial (PS) cells and in L15 medium with 5% FBS. Briefly, the virus titration was done using 200 µl of L15 medium (5% FBS, 1% penicillin/streptomycin, and 1% glutamine) in a 24-well plate. Then, a serial dilution of each virus stock from ZIKV<sup>BR</sup>, ZIKV<sup>AF</sup> and YFV-17D in L15 medium was performed, from 10<sup>-1</sup> to 10<sup>-11</sup>. Then, 200 µl of each dilution was added in each 24-well plate. After this, 1 × 10<sup>6</sup> PS cells were seeded in each well of a 24-well plate for at least 3 h at 37°C to allow virus adsorption and PS cell adherence. Later, each well was overlaid with complete carboxymethyl cellulose (CMC) medium (0.6% in L15 supplemented with 3% FBS). After 5 days of incubation at 37°C, the plaque visualization was made using blue-black staining solution. The most appropriate viral dilution was estimated to determine the amount of infected cells visible (PFU ml<sup>-1</sup>). For ZIKV<sup>BR</sup>, the first C6/36 subculture had a titre of 6 × 10<sup>8</sup>. The following subcultures had titres of 7.5 × 10<sup>6</sup> (T2) and 4 × 10<sup>12</sup> (T3). All the subculture aliquots were stored in cryovials and maintained in liquid nitrogen or were distributed to the ZIKV São Paulo task force.

**In vivo infection.** Pregnant mice, 6–8 weeks of age, C57BL/6 or SJL (JAX), were infected intravenously with 200 µl of ZIKV<sup>BR</sup>-infected C6/36 cell supernatant containing 1 × 10<sup>3</sup>, 4 × 10<sup>10</sup> or 1 × 10<sup>12</sup> PFU ml<sup>-1</sup> of virus on day 10–13 of gestation. The animals were observed daily. All the experiments were performed with the approval of the Institute of Biomedical Sciences Ethics Committee protocol number 05/2016.

**Real-time PCR.** RNA was extracted from each sample (cells, supernatant of cell culture or mouse tissue) using the QIAamp UltraSens Virus Kit (Qiagen) or TRIzol reagent (Invitrogen). All RNA pellets were resuspended in 30 µl of RNase-free distilled water, quantified using a NanoDrop spectrophotometer (NanoDrop Technologies) and stored at -80°C. The set of primers/probes specific for ZIKV were synthesized by Sigma Life Science, with 5'-FAM as the reporter dye for the probe. The set of primers/probes ZIKV 835, ZIKV 911c and ZIKV 860-FAM were previously described<sup>2</sup>. The real-time reaction was performed with 10 µl of each sample and 10 µl of the AgPath-IDTM One-Step RT-PCR reagents (Applied Biosystems). The amplification was done in an Applied Biosystems 7500 real-time PCR system, and involved activation at 45°C for 15 min, 95°C for 15 min followed by 40 amplification cycles of 95°C for 15 s, 60°C for 15 s, and 72°C for 30 s. The real-time data were analysed using SDS software from Applied Biosystems. For the detection and quantification of viral RNA, the real-time PCR of each sample was compared with threshold cycle (C<sub>t</sub>) value with a ZIKV plasmid standard curve, which was obtained carrying out serial dilutions of a clone of the envelope gene of an isolate from the 2007 Yap Island outbreak (provided by the Institute Pasteur in Dakar, Senegal). For the detection and quantification of YFV RNA, a YFV-specific real-time assay was applied. The fold changes of gene expression were calculated in comparison to the values for the YFV controls<sup>31</sup>. The positive PCR control C<sub>t</sub> value was used to normalize gene expression and determine fold changes during the 96 hours post-infection. The RPLA27 gene (60S ribosomal protein L27) was used as endogenous control for the PCR reactions<sup>30</sup>. For the TAM receptors detection,

NPCs or neurons from two donors were or were not infected with ZIKV<sup>BR</sup>, at 10 MOI for 48 h and were submitted to standard TRIzol (Invitrogen) protocol for RNA extraction. Total RNA was quantified using a NanoDrop spectrophotometer and 2 µg was used for further cDNA synthesis using the superscript III reverse-transcription polymerase (Invitrogen) according to manufacturer's protocol. qPCR was performed using Taqman probes (Extended Data Table 1) and the QuantStudio 12K Flex Real-Time PCR System (Applied Biosystems). For normalization *ACTB* was used as a housekeeping gene.

**Cell death pathway analysis.** One microgram of total RNA from brains of 4 pups, 2 pooled mock and 2 pooled ZIKV<sup>BR</sup>-infected from SJL mothers were submitted to gene expression analysis for cell death target genes using the RT2 Profiler PCR Array Mouse Cell Death PathwayFinder (cat. no. PAMM-212ZA- Qiagen) according to the manufacturer's protocols. qPCR was performed in the QuantStudio 12K Flex Real-Time PCR System (Applied Biosystems). To evaluate gene expression, we established a fold change threshold of at least twofold up- or downregulation compared to mock infected samples. Statistical analysis was performed using the RT2 profiler RT-PCR array data analysis software v3.5.

**NPCs, neurons, neurospheres and organoids.** We used three human and two chimpanzee iPSC clones that were previously characterized in the Beltrão-Braga and Muotri laboratories<sup>32–35</sup> plus H9 human embryonic stem cells (hESC) for all the experiments using pluripotent stem cells. All the cell lines tested negative for mycoplasma contamination. Briefly, high passages of iPSC/hESC colonies on feeder-free plates were maintained for 5 days with mTSeR media (Stem Cell Technologies). On the fifth day, the medium was changed to N2 media (DMEM/F12 medium supplemented with 1X N2 supplement (Invitrogen) and the dual SMAD inhibitors, 1 µM dorsomorphin (Tocris) and 1 µM SB431542 (Stemgent), for 48 h. Further, the colonies were detached from the plate and cultured in suspension as embryoid bodies (EBs) for 5 days at 90 r.p.m. in N2 media with the dual SMAD inhibitors. The EBs were plated on matrigel-coated plates with NBF media composed of the following: DMEM/F12 media supplemented with 0.5X N2, 0.5X B27 supplement (Gibco), 20 ng ml<sup>-1</sup> of FGF2 and 1% penicillin/streptomycin. The emerged rosettes containing the NPCs were manually picked, dissociated and plated in a double-coated plate with poly-ornithine (10 µg ml<sup>-1</sup>, Sigma-Aldrich) and laminin (2.5 µg ml<sup>-1</sup>, Gibco). The NPC population was expanded using NBF media. The neuronal differentiation induction protocol consisted of treating the confluent NPC plate with 10 µM ROCK inhibitor for 48 h (Y-27632, Calbiochem) in the absence of FGF in the media, with regular media changes every 3 or 4 days. Neurons were considered completely differentiated and ready to experiments after 28 days. To produce neurospheres, NPC were scrapped from the plates and submitted to continuous shaking for 5–7 days at 90 r.p.m. in NBF media. Cerebral organoids were generated as previously described<sup>25,36</sup>. All experiments were performed with the approval of the Institute of Biomedical Sciences Ethics Committee protocol number 1001.

**In vitro infection.** NPCs, neurons, neurospheres and organoids were infected with ZIKV<sup>BR</sup>, ZIKV<sup>AF</sup>, YFV and mock (culture supernatant from uninfected C6/36 cells). NPCs were seeded in plates in 24-well plates and after 24 h viral samples were diluted to the desired MOI (0.1; 1 or 10) and added to the cells. For viral adsorption, cells in monolayer were incubated for 1 h at 4°C with gentle agitation every 10 min. Next, the inoculum was removed and cells were washed once with PBS (USB Corporation). Culture medium was added to each well, and cells were incubated at 37°C and 5% CO<sub>2</sub> for the duration of the experiment. For neurospheres, NPCs were kept in constant shaking. For neuronal infection, NPCs were previously differentiated for 28 days and then neurons were infected with the desired MOI. For organoids, the number of cells available for infection was estimated to be 2.5 × 10<sup>4</sup> cells, as calculated by dividing the average surface area of a typical organoid by the average area of a typical cell (that is, a fibroblast). This calculus was used to estimate the desired 0.1 MOI. For mock controls, the same volume of supernatant was added to each experiment, and the same procedures were followed.

**Immunofluorescence.** Cells were fixed using paraformaldehyde, 4% in PBS, for 15 min at room temperature. After washing, the cells were permeabilized with 0.1% Triton X-100 (Promega) diluted in PBS for 15 min. After blocking with 2% of BSA (Sigma-Aldrich) for 4 h, primary antibodies directed against the following were added: anti-ZIKV (polyclonal mouse, Institute Pasteur in Dakar, 1:80), anti-Flavivirus D1-4G2-4-15 (polyclonal mouse, Millipore, 1:100), 1:50, anti-MAP2 (chicken, Abcam ab5392, 1:200), anti-cleaved-caspase-3 (rabbit, Cell Signaling #9661, 1:400), anti-Sox2 (mouse, Abcam ab97959), anti-GFAP (rabbit, Abcam, 1:500) and anti-Mushashi1 (rabbit, Abcam, 1:1000) (Extended Data Table 2). The cells were incubated overnight at 4°C. Secondary antibodies were added for a one-hour incubation at room temperature: anti-mouse Alexa Fluor 488, anti-chicken Alexa Fluor 647, anti-rat Alexa Fluor 555 and anti-rabbit Alexa Fluor 555 (Invitrogen). The nuclei were stained with DAPI (Invitrogen, 1:10,000) diluted in a PBS 1 × solution for 5 min and mounted with DPX (Sigma). Images were acquired with Nikon Eclipse 80i. Analysis of data was performed using software NIS Elements 3.22 (Tokyo, Japan).

**Cerebral organoids analyses.** Human and chimp organoids were infected with an MOI of 0.1 and analysed after 24 and 96 h post-infection. Organoids were cryo-sectioned at 20 µm. Immunofluorescence was performed after blocking sections in a solution with 0.1% Triton and 3% BSA (Gemini) for 1 h at room temperature. The primary antibodies were diluted in a solution with 0.1% Triton and 3% BSA, and the sections were incubated with following antibodies: anti-ZIKV, anti-Flavivirus, anti-MAP2, anti-cleaved-caspase-3 and anti-Sox2, all mentioned above, and anti-PAX6 (rabbit, Covance PRB-278P, 1:100), anti-TBR1 (rabbit, Abcam ab31940, 1:300), CTIP2 (rat, Abcam ab18465, 1:100) and Ki67 (rabbit, Abcam ab15580, 1:100). The sections were blocked with 0.1% Triton (Sigma-Aldrich) and 3% BSA for 30 min at room temperature and the secondary antibodies previously diluted, the same mentioned above, were added. The nuclei were stained with DAPI, as mentioned above and slides were mounted with DPX (Sigma-Aldrich).

**Transmission electron microscopy.** Cell pellet was fixed using a 3% glutaraldehyde solution (Merck) at 4 °C for 2 h, rinsed in three changes of PBS for 1 h, and incubated for 16 h at 4 °C. The next day, post-fixation was performed with 1% of osmium tetroxide for 30 min at room temperature. Dehydration was carried out gradually with a series of ethanol concentrations: 70%, 95% and 100%. Sample was taken through two changes of propylene oxide and placed at a 1:1 ratio with embedding medium for 1 h in a rotary mixer followed by 100% embedding medium at room temperature for 24 h. Fresh embedding medium was placed overnight at 37 °C and polymerized in oven to 24 h. Ultrathin sections were cut and stained with uranyl acetate and lead citrate. The cells were visualized with a transmission electron microscope (JEOL, JEM 1011, Peabody, Massachusetts, USA). All experimental analyses were performed blinded to the treatment.

**Flow cytometry.** Cells were infected with an MOI of 10 or 1, prepared using supernatants from infected C6/36, and equal volume of mock. Cellular infection occurred at 4 °C for 1 h with cell homogenization every 10 min. After that, the cells were washed once and then maintained at 37 °C in CO<sub>2</sub> incubators with medium, as described before. After 24, 48, 72 and 96 h post-infection the cells were harvested and then submitted to a staining protocol for annexin V and propidium iodide (PI) (BD Biosciences). The cells were washed twice with PBS and were harvested with 200 µl of trypsin 0.25% (LGC) for 10 min at 37 °C. Next, the cell suspensions were washed in DMEM with 10% of FBS and centrifuged for 5 min at 450g and 4 °C. The cells were then resuspended in 20 µl of annexin V binding buffer in 96-well round-bottom plates and with 1 µl of FITC-annexin V + 1 µl of PI and then incubated at room temperature for 15 min protected from light. After incubation period, the samples were added to 80 µl of binding buffer and acquired in the BD FACS Accuri C6 (BD Biosciences) flow cytometer.

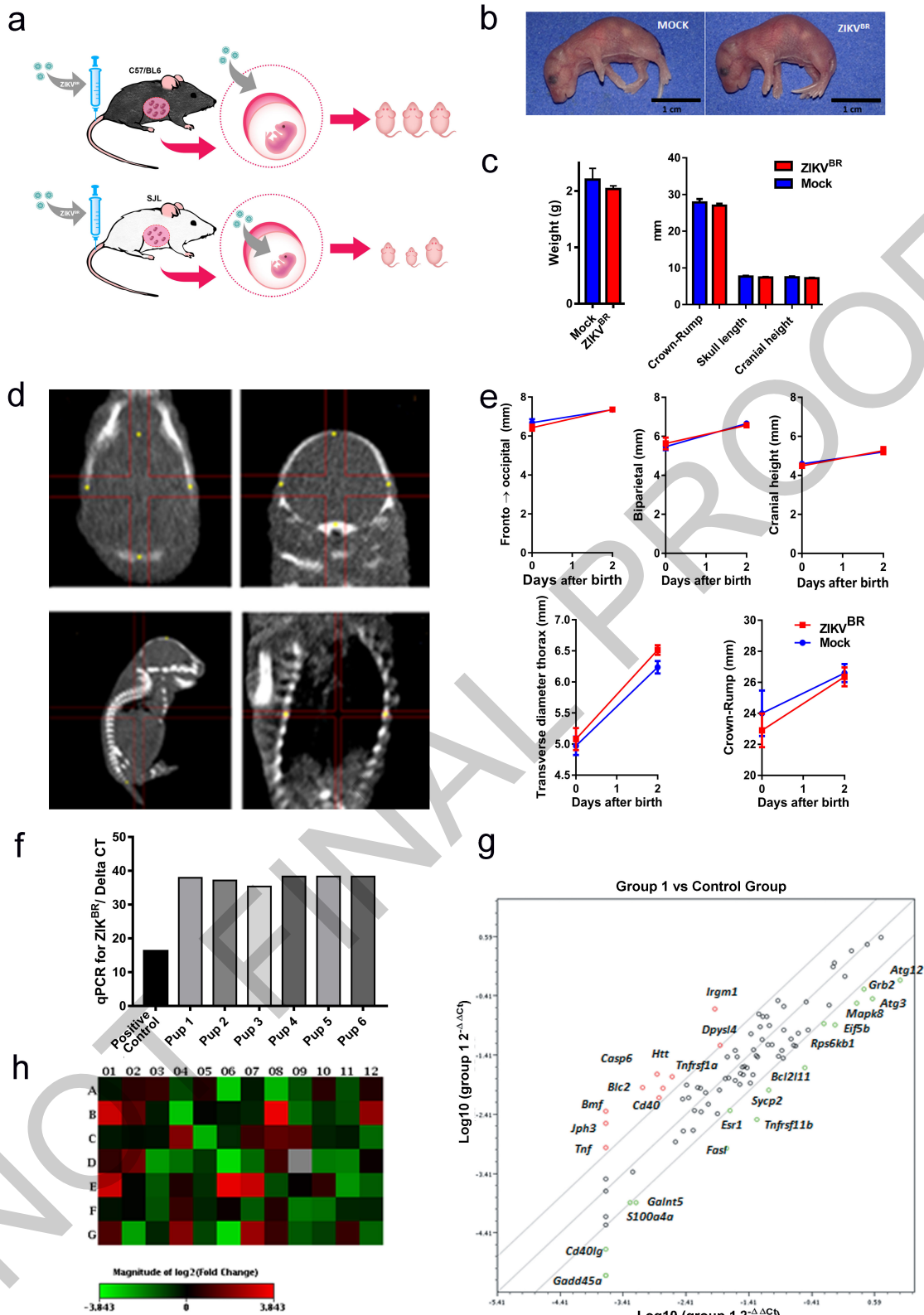
**Computed tomography.** Mice were properly anaesthetized with isoflurane and immobilized on their right side on the bed with a piece of gauze and positioned with the whole body in the field of view (FOV). CT images were acquired using small animal imaging equipment (Triumph Trimodality Gamma Medica Ideas) with 45 kVp, 0.4 mA and 2.13 min of X-ray exposition (512 projections over 360°).

The images were reconstructed using the FBP (filtered back projection) algorithm, a matrix of 512 × 512, smoothing filter and a pixel size of 92–117 µm (according to the animal size). Experienced evaluators, who were blinded to the animal group assignments, performed images analyses using the AMIDE 1.0.4, General Public License software. Fiducial marks were added to measure the distance between points considering, visually, the larger axis of the brain in the sagittal plane for posterior–anterior and superior–inferior distances and the coronal plane for the lateral right–left. For measuring whole-body length, the distance between the superior point of the brain to the first vertebra of the tail was used. The thorax measure was made from the third rib (right–left) in the coronal plane to the spinal cord level. The results are presented in mm.

**Histologic processing.** Tissue histology was performed using a dehydrating protocol with two alcohol baths of 95% (the first for 1 h and 15 min and the second for 30 min), three absolute alcohol baths (1 h for the first followed by 3 h for the second and 2 h for the third), followed by clarification with three baths of xylene (the first for 30 min and the next two for 1 h each). Finally, the material was added to three paraffin baths (the first for 30 min and the second two for 1 h each). Then the material was immersed in paraffin and cut with a microtome to a thickness of 5 mm. The deparaffinization protocol consisted of three xylol baths heated in an oven for 30 min each, two baths of absolute alcohol for 2 min each, two baths of 95% alcohol lasting 2 min each, an alcohol in water bath (85%) for 2 min and the last bath in 70% alcohol for 2 min. The haematoxylin and eosin staining protocol began with two quick baths in running water, followed by a 2-min bath in distilled water, a 2-min bath in haematoxylin, a 5-min bath in running water, a 1-min bath in eosin, followed by 1-min in a fast-flowing water bath, two baths of 95% ethanol for 2 min each, two baths of absolute ethanol for 2 min each, ending with three baths in xylene for 2 min each. Slides were mounted using Permount (Sigma-Aldrich) and analysed on multiple coronal slices in glass slides using light microscopy (Olympus BX40, ZEISS KS400) on a genotype-blinded fashion.

31. Stock, N. K. *et al.* Biological and phylogenetic characteristics of yellow fever virus lineages from West Africa. *J. Virol.* **87**, 2895–2907 (2013).
32. de Jonge, H. J. *et al.* Evidence based selection of housekeeping genes. *PLoS One* **2**, e898 (2007).
33. Marchetto, M. C. *et al.* A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. *Cell* **143**, 527–539 (2010).
34. Marchetto, M. C. *et al.* Differential L1 regulation in pluripotent stem cells of humans and apes. *Nature* **503**, 525–529 (2013).
35. Beltrao-Braga, P. I. *et al.* Feeder-free derivation of induced pluripotent stem cells from human immature dental pulp stem cells. *Cell Transplant.* <http://dx.doi.org/10.3727/096368911X566235> (2011).
36. Lancaster, M. A. & Knoblich, J. A. Organogenesis in a dish: modeling development and disease using organoid technologies. *Science* **345**, 1247125 (2014).

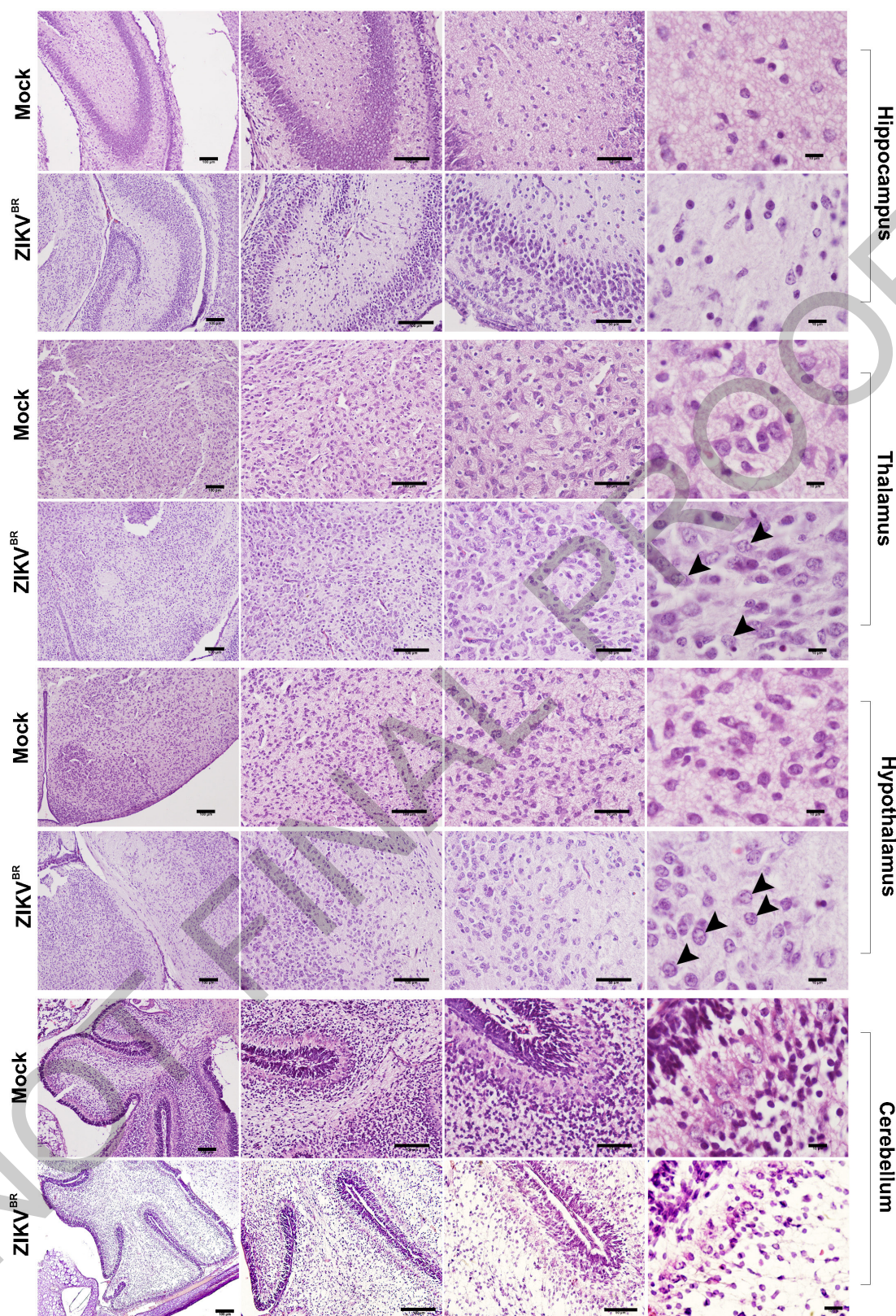




**Extended Data Figure 1 | Impact of ZIKV<sup>BR</sup> infection in the C57BL/6 and SJL mice.** **a**, Scheme for infecting mice and the follow-up analyses. Pregnant females at approximately day 10–13 of gestation were challenged with  $4 \times 10^{10}$  PFU of ZIKV<sup>BR</sup> via an intravenous route. Their pups were analysed immediately after birth for signs of malformation. **b**, A representative pup from mock-infected and the ZIKV<sup>BR</sup>-infected C57BL/6 mice. Scale bar, 1 cm. **c**, C57BL/6 pups born with no gross morphological changes or size differences compared to mock controls ( $n = 21$  pups from three separate litters, error bars, s.e.m.,  $t$ -test).

**d**, **e**, CT analysis confirmed lack of anatomical alterations ( $n = 21$  pups from three separate litters, error bars, s.e.m.,  $t$ -test). For scale, the crosses indicate 1.2 mm (top left; top right; bottom left) and 0.6 mm (bottom right). **f**, ZIKV<sup>BR</sup> RNA was not detected in the brains of six C57BL/6 pups. **g**, Cell death pathway signature revealed by qPCR gene expression in the brains of the ZIKV<sup>BR</sup>-infected SJL pups ( $n = 2$  technical replicates of pooled RNA from two pups each group; threshold = twofold). **h**, Heat map representation of misregulated genes in **g**.

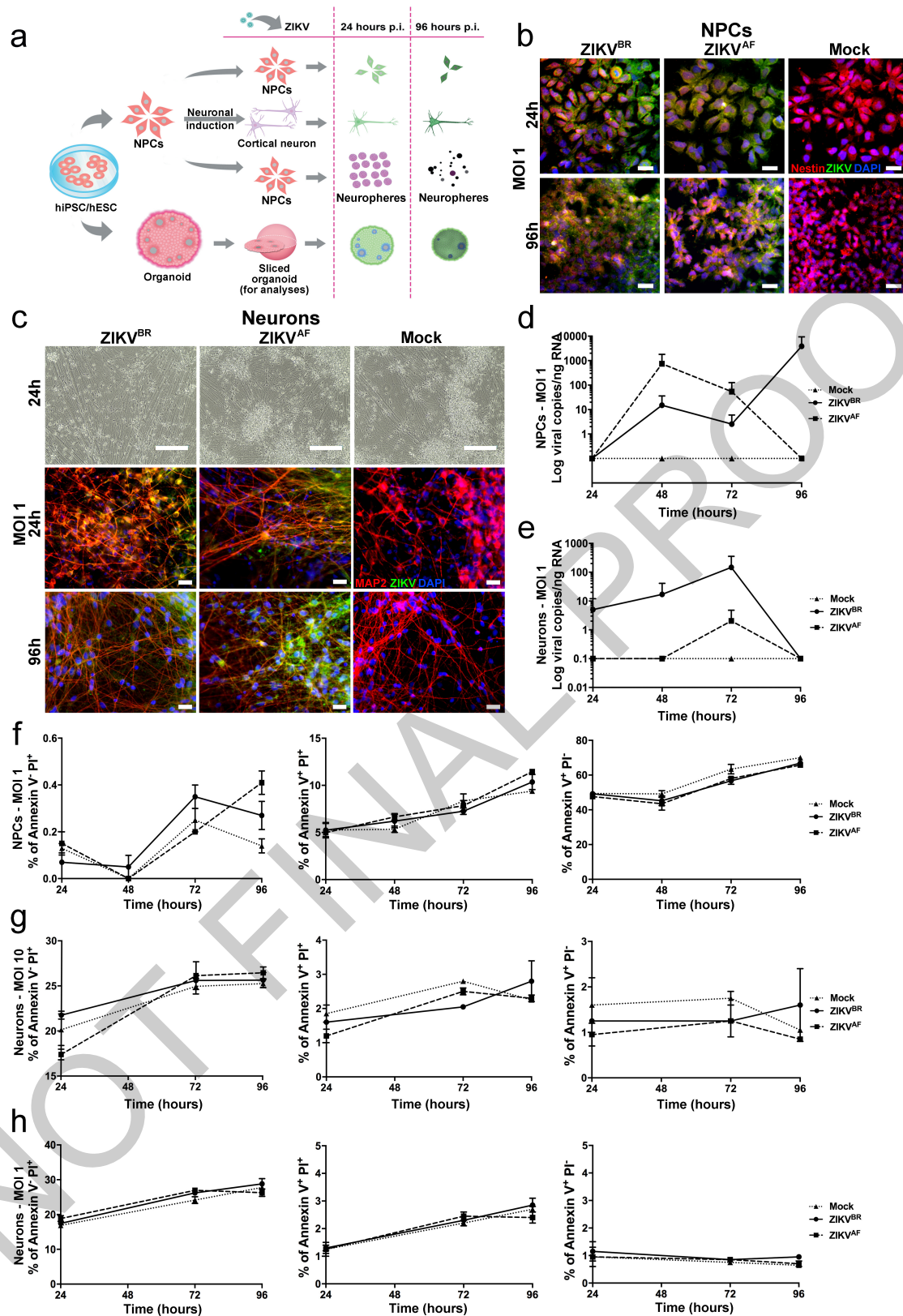




**Extended Data Figure 2 | Histopathological analysis of brains from ZIKV<sup>BR</sup>-infected SJL pups.** Morphological aspect of hippocampus, thalamus, hypothalamus and cerebellum from brains of pups born from mothers infected with ZIKV<sup>BR</sup>. Arrowheads indicate intranuclear vacuoles

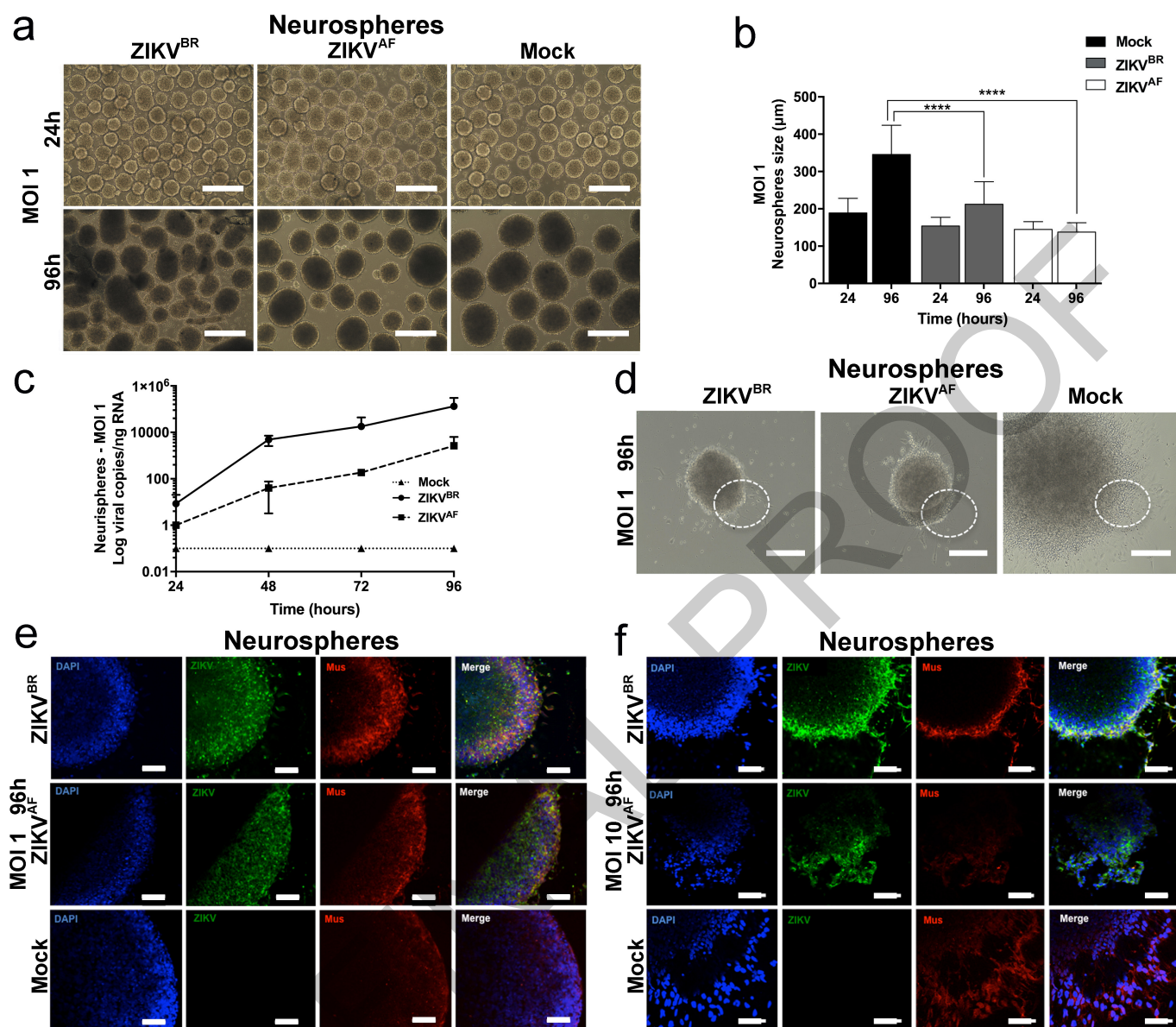
and 'empty' nuclei aspect with chromatin margination observed in thalamus and hypothalamus. Scale bar from left to right: 100  $\mu$ m, 100  $\mu$ m, 50  $\mu$ m and 10  $\mu$ m.





**Extended Data Figure 3 | Impact of ZIKV infection in human NPCs and neurons.** **a**, Scheme of the *in vitro* experiments using hPSCs. The cells were differentiated into NPCs, neurons, neurospheres and cerebral organoids to test the impact of ZIKV<sup>BR</sup> over time. **b**, Infection of NPCs with the ZIKV<sup>BR</sup> and ZIKV<sup>AF</sup> (MOI = 1) at 24 and 96 h post-infection. Scale bar, 25  $\mu$ m. **c**, Aspects of iPSC-derived human neurons after ZIKV infection (MOI = 1) at 24 and 96 h post-infection. Scale bars, 200  $\mu$ m (bright field); 25  $\mu$ m (immunofluorescence). **d**, Viral replication dynamics in human NPCs over time (MOI = 1) ( $n = 2$  technical replicates from

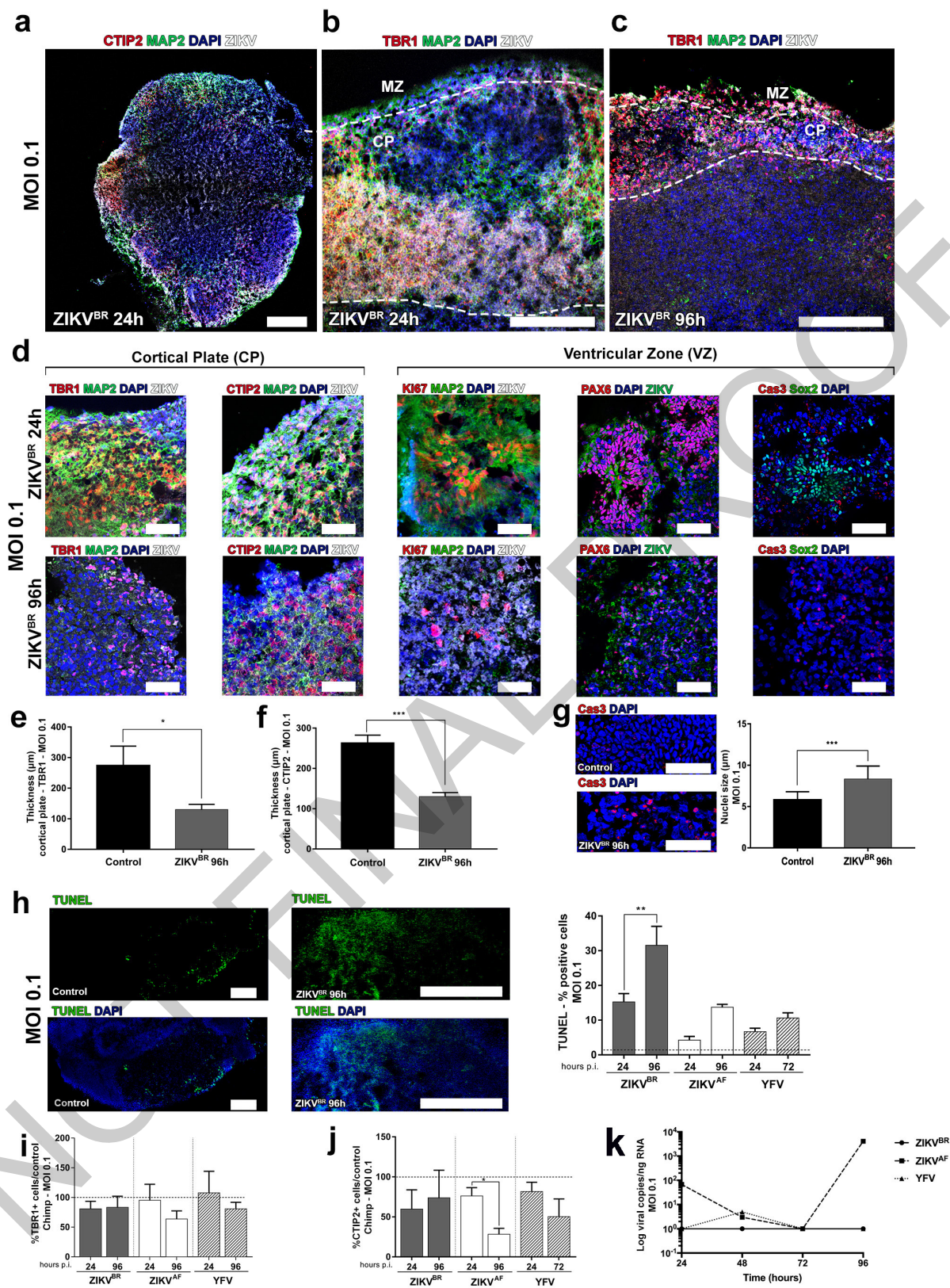
two different donors; error bars, s.e.m.). **e**, Viral replication dynamics in human neurons over time (MOI = 1) ( $n = 2$  technical replicates from two different donors; error bars, s.e.m.). **f**, Dynamics of NPCs toxicity over time after ZIKV infection (MOI = 1), indicating no significant differences among the different viruses ( $n = 2$  technical replicates from two different donors; error bars, s.e.m.). **g**, **h**, Viral replication dynamics of ZIKV in human neurons over time at MOI = 10 and MOI = 1, respectively ( $n = 2$  technical replicates from two different donors; error bars, s.e.m.; one-way ANOVA).



**Extended Data Figure 4 | Impact of ZIKV infection in human neurospheres.** **a**, Representative bright-field images of ZIKV infection (MOI = 1) at 24 and 96 h post-infection. Scale bar, 400  $\mu\text{m}$ . **b**, Alterations in neurospheres diameter over time (MOI = 1;  $n = 22$  neurospheres from two different donors for each time point in each condition; error bars, s.e.m.; ANOVA,  $*P = 0.0124$  and  $****P < 0.0001$ ). **c**, ZIKV replication dynamics in neurospheres (MOI = 1;  $n = 3$  technical replicates from two

different donors). **d**, Representative bright-field images of ZIKV infection (MOI = 1) at 96 h post-infection. Scale bars, 400  $\mu\text{m}$  and 1,000  $\mu\text{m}$  (mock). The dotted circle describes the neurospheres borders indicating the immunostained regions in **e** and **f**. **e**, **f**, Immunostaining of neurospheres infected with ZIKV at 96 h post-infection at MOI = 1 (**e**) or MOI = 10 (**f**), revealing a qualitative reduction of proliferative cell migration from Mushashi (Mus)-positive cells. Scale bars, 50  $\mu\text{m}$ .





Extended Data Figure 5 | See next page for caption.



**Extended Data Figure 5 | Human and chimp cerebral organoids infected with ZIKV.** **a**, Representative image of an entire cross-section of a cerebral human organoid infected with the ZIKV<sup>BR</sup> (MOI = 0.1, 24 h post-infection). Scale bar, 200  $\mu$ m. **b**, Detail of the surface of a human organoid infected with the ZIKV<sup>BR</sup> at 24 h post-infection (MOI = 0.1). Marginal zone (MZ) and cortical plate (CP) delineated by dotted white lines. Scale bar, 200  $\mu$ m. **c**, Detail of the surface of a human organoid infected with the ZIKV<sup>BR</sup> at 96 h post-infection (MOI = 0.1). Notice the significant tissue damage and reduction in the CP at 24 h post-infection. Scale bar, 200  $\mu$ m. **d**, A representative characterization of CP and ventricular zone (VZ) in human organoid infected with the ZIKV<sup>BR</sup> at 24 and 96 h post-infection (MOI = 0.1). Scale bar, 50  $\mu$ m. **e**, Reduction in the cortical thickness measured by the extension of TBR1-positive layer of cells in human organoids at 96 h post-infection (MOI = 0.1;  $n = 3$  replicates from three human cell lines; error bars, s.e.m.; unpaired  $t$ -test,  $*P = 0.0203$ ). **f**, Reduction in the cortical thickness measured by the extension of CTIP2-positive cells layer in human organoids at 96 h

post-infection (MOI = 0.1;  $n = 3$  replicates from three human cell lines; error bars, s.e.m.; unpaired  $t$ -test,  $***P = 0.0005$ ). **g**, Nuclear size (diameter) of cleaved-caspase-3-positive apoptotic cells in human organoids at 96 h post-infection (MOI = 0.1;  $n = 10$  organoids/slides from three human cell lines; error bars, s.e.m.; unpaired  $t$ -test,  $***P = 0.0004$ ). Scale bar, 50  $\mu$ m. **h**, Percentage of TUNEL-positive cells in relation to controls (dotted line) at 24 and 96 h post-infection (MOI = 0.1;  $n = 10$  organoids/slides from three human cell lines; error bars, s.e.m.; ANOVA,  $**P = 0.0042$ ). **i**, Percentage of TBR1-positive cells in non-primate organoids (chimp) in relation to controls (dotted line) at 24 and 96 h post-infection (MOI = 0.1;  $n = 3$  organoids from two donors; error bars, s.e.m.; ANOVA). **j**, Percentage of CTIP2-positive cells in non-primate organoids (chimp) in relation to controls (dotted line) at 24 and 96 h post-infection (MOI = 0.1;  $n = 3$  organoids from two donors; error bars, s.e.m.; ANOVA). **k**, Viral replication dynamics in chimpanzee organoids over time (MOI = 0.1;  $n = 3$  replicates from two donors; error bars, s.e.m.; ANOVA).

Extended Data Table 1 | Probes used for TAM receptors detection

Target	Species	Code
Tyro-3	Human	Hs_00170723_m1
Axl	Human	Hs_01064444_m1
MertK	Human	Hs_01031973_m1
DC-Sign	Human	Hs_01588349_m1
$\beta$ -actin	Human	Hs_01060665-m1

Extended Data Table 2 | Antibodies used in this study, related to experimental procedures

Cell type/tissue	Antigen	Host	Supplier	Cat number	Dilution
Zika virus	ZIKV	Mouse polyclonal	Institute Pasteur in Dakar	MAB10216	1:50
Flavivirus	Flavivirus	Mouse polyclonal	Millipore	ab5392	1:200
Neurons	MAP2	Chicken polyclonal	Abcam	ab5392	1:200
Apoptosis	Cleaved caspase-3	Rabbit polyclonal	Cell Signaling	9661	1:400
Progenitor cell	Sox2	Mouse monoclonal	Abcam	ab79351	1:200
Astrocytes	GFAP	Rabbit polyclonal	Abcam	ab7260	1:500
Neuronal progenitor cells	Mushashi1	Rabbit polyclonal	Abcam	ab21618	1:200
Neuronal progenitor cells	PAX6	Rabbit polyclonal	Covance	PRB-278P	1:100
Pre-plate/deep-layer cortical neurons	TBR1	rabbit polyclonal	Abcam	ab31940	1:400
Deep-layer cortical neurons	CTIP2	rat monoclonal	Abcam	ab18465	1:100
Cell proliferation	Ki67	rabbit polyclonal	Abcam	ab15580	1:100
Neurons	TUJ1	mouse polyclonal	Covance	MMS-435P	1:500



# Overcoming mTOR resistance mutations with a new-generation mTOR inhibitor

Vanessa S. Rodrik-Outmezguine<sup>1\*</sup>, Masanori Okaniwa<sup>2\*</sup>, Zhan Yao<sup>1\*</sup>, Chris J. Novotny<sup>2</sup>, Claire McWhirter<sup>3</sup>, Arpitha Banaji<sup>1</sup>, Helen Won<sup>4</sup>, Wai Wong<sup>5</sup>, Mike Berger<sup>4</sup>, Elisa de Stanchina<sup>5</sup>, Derek G. Barratt<sup>3</sup>, Sabina Cosulich<sup>3</sup>, Teresa Klinowska<sup>3</sup>, Neal Rosen<sup>1,6</sup> & Kevan M. Shokat<sup>2,7</sup>

**Precision medicines exert selective pressure on tumour cells that leads to the preferential growth of resistant subpopulations, necessitating the development of next-generation therapies to treat the evolving cancer. The PIK3CA–AKT–mTOR pathway is one of the most commonly activated pathways in human cancers<sup>1</sup>, which has led to the development of small-molecule inhibitors that target various nodes in the pathway. Among these agents, first-generation mTOR inhibitors (rapalogs) have caused responses in ‘N-of-1’ cases, and second-generation mTOR kinase inhibitors (TORKi) are currently in clinical trials<sup>2–4</sup>. Here we sought to delineate the likely resistance mechanisms to existing mTOR inhibitors in human cell lines, as a guide for next-generation therapies. The mechanism of resistance to the TORKi was unusual in that intrinsic kinase activity of mTOR was increased, rather than a direct active-site mutation interfering with drug binding. Indeed, identical drug-resistant mutations have been also identified in drug-naïve patients, suggesting that tumours with activating *MTOR* mutations will be intrinsically resistant to second-generation mTOR inhibitors. We report the development of a new class of mTOR inhibitors that overcomes resistance to existing first- and second-generation inhibitors. The third-generation mTOR inhibitor exploits the unique juxtaposition of two drug-binding pockets to create a bivalent interaction that allows inhibition of these resistant mutants.**

The MCF-7 breast cancer cell line was exposed to high concentrations of either a first-generation mTORC1 inhibitor, rapamycin or a second-generation mTOR ATP competitive inhibitor, AZD8055 (a TORKi), for 3 months, until resistant colonies emerged. Deep sequencing revealed that the AZD8055-resistant (TKi-R) clones harboured an *MTOR* mutation located in the kinase domain at the M2327I position (Fig. 1a and Extended Data Fig. 1a), while two rapamycin-resistant (RR) clones contained mutations located in the FKBP12–rapamycin-binding domain (FRB domain) at positions A2034V (RR1 cells) and F2108L (RR2 cells). The clinical relevance of these mutations is supported by a case report of a patient who acquired the identical F2108L *MTOR* mutation after relapse while under treatment with everolimus<sup>5</sup> (Extended Data Table 1).

To verify that the mutations altered the efficacy of their respective drugs and were not simply passenger mutations, we analysed the phosphorylation of effectors downstream of mTOR in several cellular systems. In the RR cells, phosphorylation of the normally rapamycin-sensitive sites on S6K (T389) and S6 (S240/244 and S235/236) were unaffected even at high rapalog concentrations (100 nM) (Fig. 1b and Extended Data Fig. 1b). Phosphorylation of the key mTOR effector 4EBP1 is normally unaffected by rapamycin but strongly reduced by

TORKi<sup>6–8</sup>. In the TKi-R cells, however, 4EBP1 phosphorylation was significantly less sensitive to a variety of TORKi (Fig. 1c and Extended Data Fig. 1c, d). Consistent with this weakened signalling inhibition, the RR and TKi-R clones were significantly less sensitive to their respective drugs in a 72 h proliferation assay when compared to the parental line (Fig. 1d, e and Supplementary Table 1). To determine if the RR and TKi-R *MTOR* mutations were directly responsible for the drug-resistance phenotype, each mutant was expressed in another model, MDA-MB-468 cells, which confirmed that the *MTOR* mutations are sufficient to promote dominant resistance (Extended Data Fig. 2a–d).

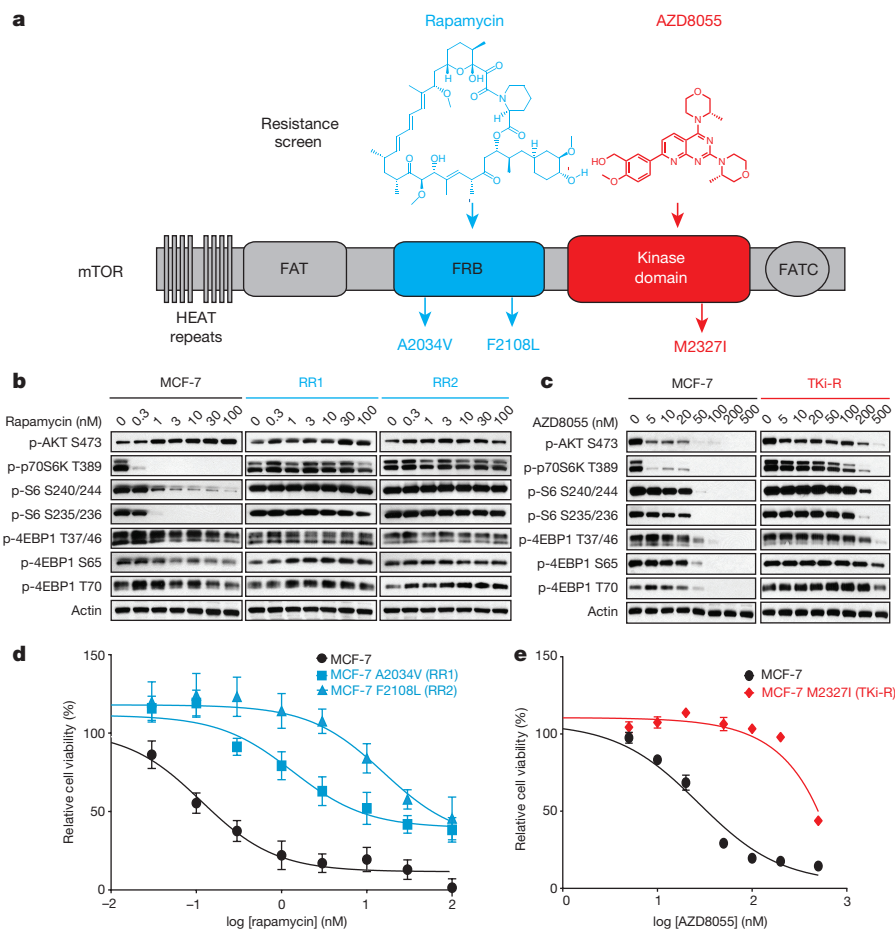
FRB domain mutations have been found in untreated patients (Extended Data Table 2), and previous random mutagenesis screens in yeast have shown that single amino acid changes in the mTOR FRB domain confer rapamycin resistance<sup>9–12</sup>. The RR mutants identified in this screen exhibit a similar mechanism of resistance by disrupting the interaction of mTOR with the FKBP12–rapamycin complex in cells and *in vitro* (Fig. 2a, b).

In contrast to the FRB-domain mutations found in RR cells, which line the rapalog/FKBP-binding pocket, analysis of the recently solved structure of the mTOR kinase domain in complex with the TORKi, PP242 (Protein Data Bank (PDB), 4JT5)<sup>13</sup>, revealed that M2327 is >15 Å away from the inhibitor, suggesting either an allosteric mechanism of reduced TORKi affinity or that this mutation causes resistance through a mechanism that does not involve reduced drug binding. Indeed, both wild-type and M2327I mTOR bind AZD8055 with similar affinities (Fig. 2c). We asked whether the M2327I mutation in the mTOR kinase domain altered the kinetic properties of the kinase. As shown in Fig. 2d, the M2327I mutant has a threefold increase in mTOR kinase activity compared with the wild-type and RR mutants. This is consistent with the higher phosphorylated (p)-S6K (T389), p-AKT (S473) and p-4EBP1 S65 basal levels observed in these cells (Extended Data Fig. 1d).

The emergence of a hyperactive *MTOR* kinase domain mutation (M2327I) that could theoretically confer a growth advantage led us to wonder if similar mutations might pre-exist in drug-naïve patient tumours. Indeed, the precise M2327I mutation as well as other *MTOR* kinase domain mutations have been identified in five untreated patients (Extended Data Tables 1 and 3)<sup>14,15</sup>. To determine if additional *MTOR* kinase domain mutants were also hyperactive and insensitive to TORKi, various *MTOR* kinase domain mutations that occur in patients were inducibly expressed in MDA-MB-468 cells and tested for sensitivity to the TORKi AZD8055 and MLN0128 (Extended Data Fig. 2d, e). The concentrations of drug required to inhibit mTORC1 and mTORC2 substrates in these cells were 3- to 30-fold higher than

<sup>1</sup>Program in Molecular Pharmacology, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. <sup>2</sup>Howard Hughes Medical Institute and Department of Cellular and Molecular Pharmacology, University of California San Francisco, San Francisco, California 94158, USA. <sup>3</sup>AstraZeneca, Alderley Park, Macclesfield, Cheshire SK10 4TG, UK. <sup>4</sup>Human Oncology and Pathogenesis Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. <sup>5</sup>Anti-Tumor Assessment Core, Memorial Sloan Kettering Cancer Center, New York, New York 10065, USA. <sup>6</sup>Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, New York 10065, USA. <sup>7</sup>Department of Chemistry, University of California Berkeley, Berkeley, California 94720, USA.

\*These authors contributed equally to this work.



**Figure 1 | Single amino acid mutation accounts for acquired resistance to mTOR inhibitors.** **a**, Graphic representation of mTOR domains and site mutagenesis isolated in rapamycin- and AZD8055-resistant cells. **b**, **c**, The effects of rapamycin (**b**) or AZD8055 (**c**) on mTOR signalling was assessed in MCF-7, RR1 and RR2 cells (or in TKi-R cells (**c**)) by immunoblotting 4 h after treatment. For gel source data, see Supplementary Fig. 1.

those required in wild-type cells, although not all substrates show precisely the same dose response.

These data suggest that the hyperactivation of mTOR kinase by single amino acid mutations found in drug-naïve patients can reduce the sensitivity to ATP-competitive mTOR inhibitors in cells. These findings highlight the need for a new class of mTOR inhibitor capable of targeting both drug-naïve (pre-existing) *MTOR* mutant-driven cancers, as well as emergent resistant mutations.

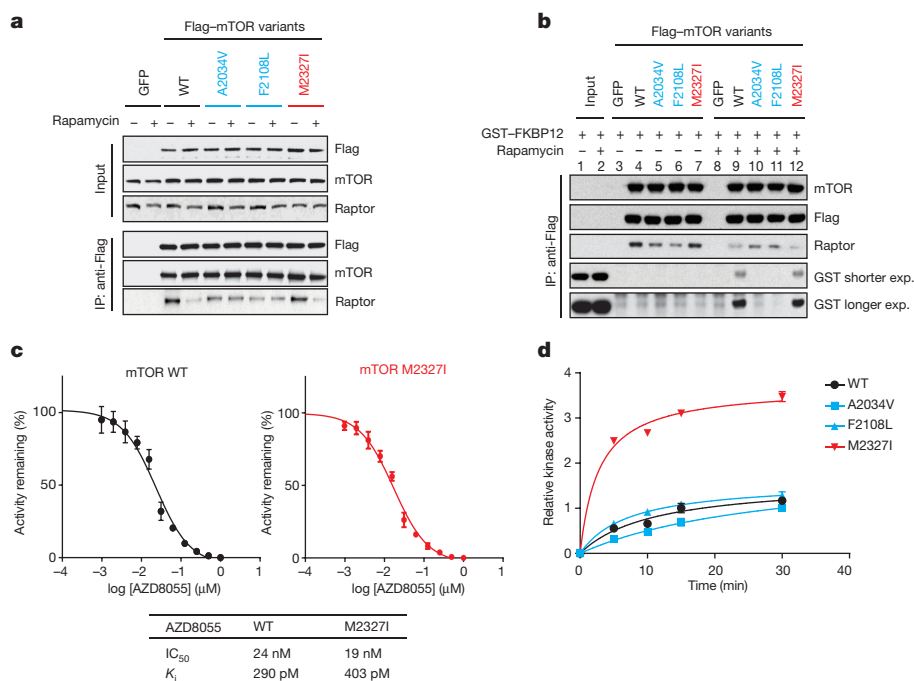
We developed a molecular model of mTOR in complex with rapamycin–FKBP12 using the FRB domain as the common domain in two available mTOR crystal structures (PDB, 1FAP and 4JT5) (Fig. 3a). This model revealed the juxtaposition of the rapamycin- and TORKi-binding sites and suggested an avidity-based approach to overcome drug-resistant mutations in either the FRB or the kinase domain. A bivalent mTOR inhibitor consisting of a rapamycin–FRB-binding element appropriately linked to a TORKi would be expected to inhibit the RR class of FRB-domain mutants because the TORKi-binding site would provide high-affinity recognition. For the TKi-R class of kinase domain mutations, a bivalent inhibitor would be predicted to be similarly potent by virtue of an intact rapamycin-binding site. We reasoned that binding at one site would position the second half of the ligand in close proximity for binding to the second site, thus overcoming point mutations that diminish drug binding (as found in RR cells) or that hyperactivate the kinase (as found in TKi-R cells)<sup>16</sup>. To develop a new bivalent class of mTOR inhibitors, we required a non-perturbing, strain-free linker between rapamycin and a TORKi,

**d**, **e**, Dose-dependent cell growth inhibition curves of MCF-7 and rapamycin-resistant MCF-7 A2034V (RR1) and MCF-7 F2108L (RR2) cells treated with rapamycin at day 3 (**d**) or MCF-7 and AZD8055-resistant MCF-7 M2327I (TKi-R) cells treated with AZD8055 (**e**). Each dot and error bar on the curves represents mean  $\pm$  standard deviation (s.d.) ( $n = 8$ ). All experiments were repeated at least three times.

such that the resulting inhibitor can simultaneously bind to both sites. Analysis of our mTOR–rapamycin–FKBP12 model revealed that the hydroxyl group at the C40 position of rapamycin is exposed to solvent and is oriented towards the ATP-binding site of mTOR (Fig. 3a). Analysis of the TORKi (PP242)-bound structure (PDB, 4JT5) revealed that the N-1 position of the pyrazole ring is oriented towards rapamycin and exposed to solvent (Extended Data Fig. 3a). We selected MLN0128 as the TORKi as it is a highly selective<sup>17</sup> structural analogue of PP242, and is currently in clinical trials.

To determine the optimum linker length between the chosen sites, we used the molecular modelling program Molecular Operating Environment<sup>18</sup> to evaluate the potential energy of a methylene-based cross-linker with lengths from 10 to 40 heavy atoms. This analysis revealed that 27 atoms would be the minimal length required to span the two ligand-binding sites (Extended Data Fig. 3b). We incorporated a polyethylene glycol unit of varying lengths and used the azide-alkyne cycloaddition reaction to synthesize RapaLink-1, -2 and -3 (Fig. 3b, Extended Data Fig. 3c and Supplementary Methods). Our modelling suggested that RapaLink-3, with an 11-heavy-atom linker would be too short to allow optimal binding to both sites simultaneously, while RapaLink-1 and -2, which contain 39- and 36-heavy-atom linkers, respectively, would allow simultaneous bivalent binding to the mTOR–FKBP12 complex.

Cells were treated with increasing concentrations of either RapaLink-1, -2 or -3, and the effects on mTOR signalling were assessed by western blotting. We observed that both RapaLink-1 and -2

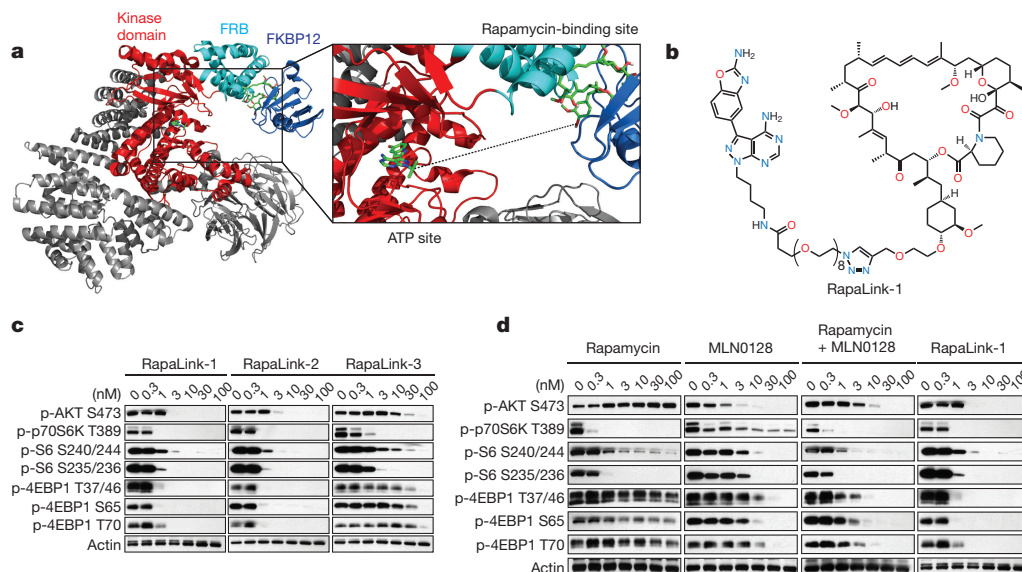


**Figure 2 | Non-overlapping mechanisms of resistance mediated by mTOR mutations.** **a**, mTOR-Flag wild-type (WT) and variants were transfected into 293H cells. Cells were treated with rapamycin and lysates were immunoprecipitated (IP) with an anti-Flag antibody. mTORC1 complex formation was assessed by immunoblotting. **b**, 293H cells were transfected and complex isolated as described in **a**, and an *in vitro* competition assay was performed followed by immunoblotting. Shorter and longer exposure (exp.) are shown. For gel source data, see Supplementary Fig. 2. **c**, Varying concentrations of AZD8055 were tested *in vitro* on wild-type and M2327I mTOR followed by a kinase reaction

(see Methods). The half-maximum inhibitory concentration (IC<sub>50</sub>) values were determined by fitting to a standard four-parameter logistic using GraphPad Prism v.5. The diagram shows the mean of  $n = 3$  data. The error bars represent the s.d. between experiments. **d**, 293H cells were transfected and the complex was isolated as described in **a**. An *in vitro* kinase assay was performed and the level of p-AKT (S473) was determined by immunoblotting. Symbols on each curve represent the relative p-AKT at different time points. The kinase activity curves were generated using GraphPad Prism v.6 after densitometry analysis was performed. All experiments were repeated at least three times.

inhibited the phosphorylation of both mTORC1 and mTORC2 targets at doses between 1 and 3 nM (Fig. 3c). However, RapaLink-3, which contains the shortest linker, showed diminished potency against the phosphorylation of 4EBP1 (T37/46/70 and S65) and AKT (S473) while still inhibiting p-S6 (S240/244 and S235/236). This is consistent with

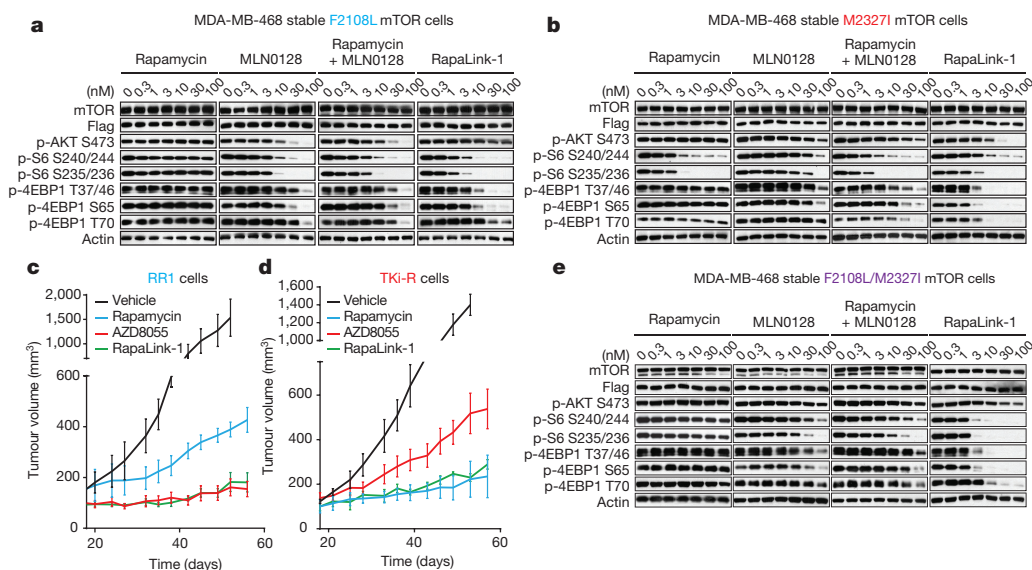
the prediction that a longer linker is necessary to allow simultaneous binding to both drug sites and indicates that rapamycin binding is dominant over MLN0128 binding due to the preferential inhibition of p-S6 over p-4EBP1. Consistent with its strong signalling inhibition (Fig. 3d), RapaLink-1 potently inhibited the growth of MCF-7 cells at



**Figure 3 | RapaLink-1 is a potent mTOR inhibitor.** **a**, Molecular model constructed by two available co-crystal structures, mTOR catalytic-domain-bearing TORKi PP242 (PDB, 4JT5) and mTOR FRB-domain-rapamycin-FKBP12 (PDB, 1FAP). Dotted line represents a guide line for the linker design of bivalent mTOR inhibitors. **b**, RapaLink-1 structure is displayed. **c**, **d**, MCF-7 cells were treated with RapaLink-1, -2 and -3 (**c**)

or with rapamycin, MLN0128, or a combination of rapamycin and MLN0128 or RapaLink-1 (**d**) for 4 h followed by immunoblotting. The rapamycin panel is the same as that shown in Fig. 1b and the RapaLink-1 panel is the same as that shown in **c**. All cellular experiments were repeated three times. For gel source data, see Supplementary Fig. 3.





**Figure 4 | RapaLink-1 reverses resistance due to mTOR FRB and kinase domain mutations.** **a, b, e,** MDA-MB-468 cells inducibly expressing mTOR F2108L (**a**) or M2327I (**b**) or F2108L/M2327I mTOR double mutant (**e**) were treated as in Fig. 3d, followed by immunoblotting. For gel source data, see Supplementary Figs 4, 5 and 6. All experiments were repeated at least three times. **c, d,** Mice bearing RR1 (**c**) or TKI-R (**d**)

xenograft tumours ( $n = 5$  for each group) were randomized to four different groups: (1) vehicle (Monday (M), Wednesday (W), Friday (F)); (2) rapamycin ( $10 \text{ mg kg}^{-1}$ ; M, W, F); (3) AZD8055 ( $75 \text{ mg kg}^{-1}$ ; M, W, F); and (4) RapaLink-1 ( $1.5 \text{ mg kg}^{-1}$ ; weekly). Tumour size was measured by calliper twice per week. The results were reported as tumour volume ( $\text{mm}^3$ )  $\pm$  s.d.

levels comparable to rapamycin or a combination of rapamycin with MLN0128 (Extended Data Figs 4a and 5).

We further tested the requirement of both halves of RapaLink-1 to simultaneously bind mTOR. First, we measured the ability of RapaLink-1 to recruit FKBP12 to mTOR by performing an *in vitro* FKBP12-binding assay; we show that RapaLink-1 is indeed able to recruit glutathione *S*-transferase (GST)-FKBP12 to wild-type mTOR (Extended Data Fig. 4b, lane 12). Moreover, we used the FKBP12 competitive ligand, FK506, to pharmacologically block RapaLink-1 from interacting with FKBP12 and thus mTOR. We observed that FK506 completely rescued the phosphorylation of mTORC1 and C2 substrates upon RapaLink-1 treatment (Extended Data Fig. 4c). Last, we isolated MCF-7 RapaLink-1-resistant cells; these cells harbour a mutation located in the mTOR FRB domain at position F2039S. As shown in Extended Data Fig. 6a, rapamycin treatment did not inhibit p-S6K (T389) and p-S6 (S240/244 and S235/236) in the mTOR F2039S cells as observed in the MCF-7 cells (Fig. 1b). Moreover, these cells displayed a decreased sensitivity to MLN0128 and a combination of rapamycin and MLN0128, as well as RapaLink-1 as compared to parental MCF-7 cells (Fig. 3d). Taken together, these data demonstrate that the binding of RapaLink-1-FKBP12 to the FRB domain is necessary for simultaneous binding to the ATP site of mTOR and therefore for RapaLink-1-dependent inhibition of mTOR signalling.

While the design of bivalent inhibitors for therapeutic use has had mixed success owing to the poor pharmaceutical properties of the hybrid molecules<sup>19</sup>, FKBP12-binding hybrids have actually been used to improve the pharmaceutical properties of small-molecule inhibitors unrelated to TORKi. These FK506-based hybrids exploit the high intracellular concentration of FKBP12, specifically in blood cells, and the high affinity of FK506 for FKBP12, to create a reservoir of drug that prolongs serum half-life<sup>20</sup>. In agreement with the improved pharmaceutical properties of previous FKBP12-binding hybrids, RapaLink-1 showed prolonged inhibition of mTOR signalling *in vitro* (Extended Data Fig. 6b, c), as well as *in vivo* after a tolerable dose of  $1.5 \text{ mg kg}^{-1}$ , which lasted for over 4 days (Extended Data Fig. 6d, e) and was able to inhibit the growth of wild-type mTOR MCF-7 xenografts as well as the current clinical mTOR inhibitors (Extended Data Fig. 6f).

To assess whether RapaLink-1 could block mTOR signalling of the F2108L mTOR and M2327I mTOR drug-resistant mutants,

MDA-MB-468 cells expressing the alleles were treated with either rapamycin, MLN0128, a combination of both drugs, or RapaLink-1. Consistent with the ability of the RapaLink-1-FKBP12 complex to bind mTOR FRB and kinase-domain mutants (Extended Data Fig. 4b, lanes 18 and 24), and increased avidity compared to rapamycin or MLN0128 (Extended Data Fig. 7a); RapaLink-1 at low doses ( $3\text{--}10 \text{ nM}$ ) was the only drug regimen capable of inhibiting mTOR signalling in both F2108L mTOR- and M2327I mTOR-expressing cells (Fig. 4a, b). Mouse xenografts of MCF-7 cells expressing the RR1 mutant A2034V mTOR showed significantly less sensitivity to rapamycin yet maintained full sensitivity to AZD8055 and RapaLink-1 treatment (Extended Data Fig. 7b and Fig. 4c). Similarly, xenografts with MCF-7 cells expressing the TKI-R mutant, M2327I mTOR, showed significantly less sensitivity to AZD8055 treatment, yet retained full sensitivity to rapamycin and RapaLink-1 (Extended Data Fig. 7c and Fig. 4d). The dosing of RapaLink-1 may be limited by toxicity, which can only be tested in the clinic. However, our preclinical data and that of others<sup>21</sup> indicate that mTOR kinase inhibitors can be given safely when administered intermittently and are more effective than daily dosing schedules.

It is reasonable to anticipate that patients bearing hyperactive *MTOR* kinase domain mutations, who originally respond to rapalogs, may eventually relapse owing to the emergence of a second FRB mutation, as previously observed<sup>5</sup>. To test whether RapaLink-1 would be an effective mTOR inhibitor in this case, MDA-MB-468 cells expressing F2108L/M2327I mTOR mutations were generated. As expected, mTOR substrates were resistant to rapamycin, MLN0128 and to a combination of both treatments in the F2108L/M2327I double-mutant cells. Yet, the signalling of these double-mutant cells remained as sensitive as the mTOR wild-type cells to RapaLink-1 treatment (Fig. 4e and Extended Data Fig. 7d).

Through exploitation of both the ATP- and the FRB-binding sites of mTOR, we have developed a new class of mTOR inhibitor that potently inhibits tumour growth and signalling in wild-type mTOR-expressing cells as well as in cells that have acquired resistance to rapalogs or ATP-competitive inhibitors, or both. Such inhibitors have been developed for G-protein-coupled receptors<sup>22</sup> (termed bitopic ligands) but have not been exploited in protein kinase inhibitor design. Interestingly, the only other bitopic kinase inhibitor we are aware of is the natural

CDK2/cyclin-A inhibitor p27. The peptidic inhibitor spans the cyclin box and extends into the ATP site of CDK2, creating a high-affinity, highly specific inhibitor<sup>23</sup>. Perhaps other allosteric sites near the ATP pocket on kinases could be similarly exploited, such as the PIF pocket<sup>24</sup>, or it might even be possible to bridge two adjacent ATP pockets in kinase complexes such as KSR–MEK<sup>25</sup>.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 29 October 2015; accepted 31 March 2016.**

**Published online 18 May 2016.**

- Vivanco, I. & Sawyers, C. L. The phosphatidylinositol 3-kinase AKT pathway in human cancer. *Nature Rev. Cancer* **2**, 489–501 (2002).
- Basu, B. *et al.* First-in-human pharmacokinetic and pharmacodynamic study of the dual m-TORC 1/2 inhibitor AZD2014. *Clin. Cancer Res.* **21**, 3412–3419 (2015).
- Iyer, G. *et al.* Genome sequencing identifies a basis for everolimus sensitivity. *Science* **338**, 221 (2012).
- Wagle, N. *et al.* Activating mTOR mutations in a patient with an extraordinary response on a phase I trial of everolimus and pazopanib. *Cancer Discov.* **4**, 546–553 (2014).
- Wagle, N. *et al.* Response and acquired resistance to everolimus in anaplastic thyroid cancer. *N. Engl. J. Med.* **371**, 1426–1433 (2014).
- Feldman, M. E. *et al.* Active-site inhibitors of mTOR target rapamycin-resistant outputs of mTORC1 and mTORC2. *PLoS Biol.* **7**, e38 (2009).
- Thoreen, C. C. *et al.* An ATP-competitive mammalian target of rapamycin inhibitor reveals rapamycin-resistant functions of mTORC1. *J. Biol. Chem.* **284**, 8023–8032 (2009).
- Dowling, R. J. O. *et al.* mTORC1-mediated cell proliferation, but not cell growth, controlled by the 4E-BPs. *Science* **328**, 1172–1176 (2010).
- Brown, E. J. *et al.* Control of p70 S6 kinase by kinase activity of FRAP *in vivo*. *Nature* **377**, 441–446 (1995).
- Chen, J., Zheng, X. F., Brown, E. J. & Schreiber, S. L. Identification of an 11-kDa FKBP12-rapamycin-binding domain within the 289-kDa FKBP12-rapamycin-associated protein and characterization of a critical serine residue. *Proc. Natl Acad. Sci. USA* **92**, 4947–4951 (1995).
- Hara, K. *et al.* Regulation of eIF-4E BP1 phosphorylation by mTOR. *J. Biol. Chem.* **272**, 26457–26463 (1997).
- Lorenz, M. C. & Heitman, J. TOR mutations confer rapamycin resistance by preventing interaction with FKBP12-rapamycin. *J. Biol. Chem.* **270**, 27531–27537 (1995).
- Yang, H. *et al.* mTOR kinase structure, mechanism and regulation. *Nature* **497**, 217–223 (2013).
- Grabner, B. C. *et al.* A diverse array of cancer-associated MTOR mutations are hyperactivating and can predict rapamycin sensitivity. *Cancer Discov.* **4**, 554–563 (2014).
- Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**, 401–404 (2012).
- Mammen, M., Choi, S. K. & Whitesides, G. M. Polyvalent interactions in biological systems: Implications for design and use of multivalent ligands and inhibitors. *Angew. Chem. Int. Ed.* **37**, 2754–2794 (1998).
- Hsieh, A. C. *et al.* The translational landscape of mTOR signalling steers cancer initiation and metastasis. *Nature* **485**, 55–61 (2012).
- Molecular Operating Environment (Chemical Computing Group Inc., Montreal, Canada, 2016).
- Szczepankiewicz, B. G. *et al.* Discovery of a potent, selective protein tyrosine phosphatase 1B inhibitor using a linked-fragment strategy. *J. Am. Chem. Soc.* **125**, 4087–4096 (2003).
- Marinac, P. S. *et al.* FK506-binding protein (FKBP) partitions a modified HIV protease inhibitor into blood cells and prolongs its lifetime *in vivo*. *Proc. Natl Acad. Sci. USA* **106**, 1336–1341 (2009).
- Patel, M. R. *et al.* A phase I study evaluating continuous and intermittent AZD2014 in combination with fulvestrant in patients with ER<sup>+</sup> advanced metastatic breast cancer (abstract). *Proc. AACR 106th Ann. Meeting* CT233.25 (AACR, 2015).
- Valant, C., Robert Lane, J., Sexton, P. M. & Christopoulos, A. The best of both worlds? Bitopic orthosteric/allosteric ligands of G protein-coupled receptors. *Annu. Rev. Pharmacol. Toxicol.* **52**, 153–178 (2012).
- Russo, A. A., Jeffrey, P. D., Patten, A. K., Massagué, J. & Pavletich, N. P. Crystal structure of the p27<sup>Kip1</sup> cyclin-dependent-kinase inhibitor bound to the cyclin A–Cdk2 complex. *Nature* **382**, 325–331 (1996).
- Wei, L. *et al.* Design and synthesis of benzazepin-2-one analogs as allosteric binders targeting the PIF pocket of PDK1. *Bioorg. Med. Chem. Lett.* **20**, 3897–3902 (2010).
- Brennan, D. F. *et al.* A Raf-induced allosteric transition of KSR stimulates phosphorylation of MEK. *Nature* **472**, 366–369 (2011).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** N.R. would like to thank the National Institutes of Health (NIH) (P01 CA094060) for funding, as well as the Breast Cancer Research Foundation grant and the National Cancer Institute Cancer Center Support grant P30 CA008748, W. H. Goodwin and A. Goodwin, the Commonwealth Foundation for Cancer Research, The Center for Experimental Therapeutics at Memorial Sloan Kettering Cancer Center, and the team up for a Cure Fund. K.M.S. would like to thank the NIH P50 AA017072, the Stand Up 2 Cancer Lung Cancer Dream Team, The Samuel Waxman Cancer Research Foundation and the Howard Hughes Medical Institute for funding. We would like to thank R. Mukherjee, S. Schwartz, J. Taunton and B. Roth for helpful comments.

**Author Contributions** V.S.R.-O., M.O., Z.Y., C.J.N., N.R. and K.M.S. conceived the project, designed and analysed the experiments, and wrote the manuscript. V.S.R.-O., M.O., Z.Y., C.J.N., C.M., A.B., W.W., D.G.B., S.C. and T.K. performed and supervised the laboratory experiments. H.W. and M.B. performed and supervised the IMPACT sequencing and analysis. E.d.S. designed and supervised the *in vivo* experiments.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to K.M.S. (kevan.shokat@ucsf.edu) or N.R. (rosenn@mskcc.org).

## METHODS

**Cell culture and reagents.** All cell lines were obtained from the American Type Culture Collection (ATCC). MCF-7 and MDA-MB-468 (ATCC catalogue numbers HTB-22 and HTB-132, respectively) breast cancer cell lines were maintained in a 1:1 mixture of DMEM:F12 medium supplemented with 4 mM glutamine, 100 units ml<sup>-1</sup> each of penicillin and streptomycin, and 10% heat-inactivated fetal bovine serum (FBS) and incubated at 37 °C in 5% CO<sub>2</sub>. The MDA-MB-468 inducible expression cells were maintained in the same medium with addition of 50 µg ml<sup>-1</sup> hygromycin and 0.2 µg ml<sup>-1</sup> puromycin. The HEK-293 cells (ATCC catalogue number CRL-1573) were maintained in DMEM medium with glutamine, antibiotics and 10% FBS. The cell lines tested negative for mycoplasma contamination. AZD8055 was obtained from AstraZeneca Pharmaceuticals, rapamycin was purchased from EMD Bioscience. RAD001, KU006, WY354, PP242 and MLN0128 were purchased from Tocris. Doxycycline was purchased from Sigma-Aldrich. Puromycin and hygromycin stock solution were purchased from Invitrogen. Drugs were dissolved in DMSO to yield 10 mM stock and stored at -20 °C.

**Selection of drug-resistant clones.** Cell lines resistant to rapamycin (RR1 and RR2) and AZD8055 (TKi-R) were generated by exposing the parental breast cancer cell line MCF-7 to a high dose of drug (500 nM of either rapamycin or AZD8055) for 3 months of continuous drug exposure (change of media every 3 days); the cells were then sent to sequencing. MCF-7-RapaLink-1-resistant cells were generated by exposing MCF-7 cells to RapaLink-1 (10 nM) for 9 months of continuous drug exposure (change of media once per week); the cells were then sent to sequencing.

**Genomic DNA sequencing.** We profiled genomic alterations in 279 key cancer-associated genes using our integrated mutation profiling of actionable cancer targets (IMPACT) assay, which utilizes solution-phase hybridization-based exon capture and massively parallel DNA sequencing. Custom oligonucleotides were designed to capture all protein-coding exons and select introns of 279 commonly implicated oncogenes, tumour suppressor genes, and members of pathways deemed actionable by targeted therapies. We prepared barcoded sequence libraries (New England Biolabs, Kapa Biosystems) for DNA from the MCF-7 parental cell line and drug-resistant subclones, and we performed exon capture on barcoded pools by hybridization (Nimblegen SeqCap). Two-hundred and fifty nanograms of genomic DNA was input for library construction. Libraries were pooled at equimolar concentrations (100 ng per library), combined with barcoded libraries from a separate project, and input to a single exon capture reaction as previously described<sup>26</sup>. DNA was subsequently sequenced on an Illumina HiSeq 2000 to generate paired-end 75-bp reads. We achieved a mean unique sequence coverage of 487 × per sample.

**Method for construction of docking modelling.** The coordinate of the crystal structure of rapamycin with FKBP12 and the FRB domain was retrieved from the PDB (accession 1FAP). The coordinate of the crystal structure of ATP-competitive mTOR inhibitor (PP242) with mTORΔN and mLST8 was retrieved from the PDB (accession 4JT5). Two co-crystal structures were aligned and amino acids and water molecules in the FRB domain of 1FAP were deleted. The coordinate of co-crystal structure unavailable ATP-competitive mTOR inhibitor (MLN0128) was manually constructed by modifying the coordinate of the co-crystal structure of its analogue (PP242). The obtained modelling containing rapamycin and MLN0128 was energy-minimized using the MMFF94x force field in Molecular Operating Environment (MOE; described later)<sup>18</sup> to provide a template structure. During the minimization procedure, the following conditions were adopted. The dielectric constant was set to 4 × *r*, where *r* is the interatomic distance. The residues, which are 9 Å away from compound, were fixed. And atomic charges for the protein and the compounds were set according to the AMBER99 and the AM1-BCC method, respectively. A crosslinker tethering rapamycin with an ATP-competitive mTOR inhibitor was manually constructed and energy-minimized using the MMFF94x force field in MOE to provide the initial conformation. The obtained initial conformation was subjected to conformation search using LowModeMD in MOE (iterative limitation was set as 30). Values of potential energies (kcal mol<sup>-1</sup>) of automatically created conformation(s) were averaged.

**Code availability.** MOE (version 2013.0801, Chemical Computing Group, Montreal, Canada), Scientific Vector Language (SVL) source.

**Cell proliferation assay.** The effect of the drug on cell proliferation was determined using a CellTiter-Glo Luminescent Cell Viability Assay kit (Promega), which is based on quantification of the cellular ATP level. Cells were plated in 96-well plates at a density of 2,000–5,000 cells (8 replicates per condition). The following day, cells were treated with a range of drug concentrations prepared by serial dilution. After 1–3 days of treatment, 100 µl of prepared reagent was added to each well. The contents of the wells were mixed on a plate shaker for 1 h, and then luminescence was measured by an Analyst AD (Molecular Devices). The relative growth was normalized to the untreated samples in each group. The growth or inhibition curves and IC<sub>50</sub> values were calculated with Graph Pad Prism v.6.

**Immunoblot analysis.** Cells were washed with PBS once, disrupted on ice for 30 min in NP-40 (50 mM Tris (pH 7.4), 1% NP-40, 150 mmol l<sup>-1</sup> NaCl, 40 mmol l<sup>-1</sup> NaF) or RIPA lysis buffer (Thermo Scientific) supplemented with protease and phosphatase inhibitors (Pierce Chemical) and cleared by centrifugation. Protein concentration was determined with BCA reagent from Pierce. Equal amounts of protein (10 to 50 µg) in cell lysates were separated by SDS-PAGE, transferred to nitrocellulose membranes (GE healthcare), immunoblotted with specific primary and secondary antibodies and detected by chemiluminescence with the ECL detection reagents from Amersham Biosciences. Antibodies for p-AKT (S473) (#4060L), p-p70S6K (T389) (#9234L), p-S6 (S240/244) (#5364L) and p-S6 (S235/236) (#4858L), p-4EBP1 (T37/46) (#9459L), p-4EBP1 (S65) (#9451L), p-4EBP1 (T70) (#9455L), β-actin (#4970S), mTOR (#2972S) and raptor (#2280S) were purchased from Cell Signaling Technology. The Flag (#F1804) antibody was purchased from Sigma. The GST antibody (#sc-138) was from Santa Cruz.

**Retrovirus-based gene-inducible expression cell system.** The mTOR genes were sub-cloned into TTIGFP-MLUEX vector harbouring tet-regulated promoter. Mutations were introduced by using the site-directed Mutagenesis Kit (Stratagene) as previously described<sup>27</sup>. The retroviruses encoding the rtTA3 or *MTOR* genes were packaged in Phoenix-AMPHO cells. The medium containing virus was filtered with 0.45 µm PVDF filters followed by incubation with the target cells for 6 h. The cells were then cultured in virus-free medium for 2 days. The cells were selected with puromycin (2 µg ml<sup>-1</sup>) or hygromycin (500 µg ml<sup>-1</sup>) for 3 days. The positive infected cell populations were further sorted with transiently expressed GFP marker after being exposed to 1 µg ml<sup>-1</sup> doxycycline and the sorted positive cells were cultured and expanded in medium without doxycycline but with antibiotics at a maintaining dose until the following assays.

**Transient transfections.** Cells were seeded at 60-mm or 100-mm plates and transfected the following day using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. The ratio between DNA and lipofectamine was 1 µg DNA to 3 µl lipofectamine.

**In vitro FKBP12 binding assay.** Cells expressing Flag-tagged wild-type or mutant mTOR were collected and lysed with 0.3% CHAPS buffer. The mTOR complexes were pulled down with anti-Flag-antibody-conjugated agarose. Then, the bead-bound complexes were incubated with recombinant FKBP12 (Fisher Scientific) (250 nM) or FKBP12 (250 nM) and rapamycin (250 nM) at 4 °C for 30 min. After incubation, the beads were washed five times with CHAPS buffer. The protein complexes were eluted with 1 × Laemmli Buffer and assayed by western blotting.

**Sequencing Sanger.** The complementary DNA was generated by messenger RNA isolated from cell pellets with SV total RNA isolation kit, SV minipreps DNA purification kit and ImProm-II Reverse Transcription System kit from Promega. The mTOR cDNA was amplified with the oligonucleotides listed in Supplementary Methods. The PCR products were subjected to gel purification and sequenced by Genewiz.

**Mutagenesis.** All the mTOR mutants were generated by QuikChange II Site-Directed Mutagenesis Kit obtained from Agilent and confirmed by Sanger sequencing.

**mTOR in vitro kinase assay.** Active mTOR kinases were expressed in 293 cells and isolated by immunoprecipitation with anti-Flag beads in 0.3% CHAPS buffer. The AKT recombinant protein was acquired from AstraZeneca Pharmaceuticals. The *in vitro* kinase assays was performed with 250 µM ATP at 37 °C for 20 min.

**In vitro kinase inhibition assay.** Concentration–response curves with a concentration range of 1,000 to 0.97 nM and twofold serial dilution were constructed by dispensing a 100 µM DMSO solubilized stock of AZD8055 into white 384-well medium-binding microplates (Greiner Bio-One) using an HP D3000 Digital Dispenser. The kinase reaction was performed as described earlier using 200 µM ATP, 1.5 µM peptide substrate and either 5 nM wild-type mTOR or 2 nM M2327I mutant mTOR. The IC<sub>50</sub> values were calculated from initial rate data before being corrected for competition with ATP using the Cheng–Prussoff equation and assuming the compound is fully ATP competitive<sup>28</sup>. The IC<sub>50</sub> values were determined by fitting to a standard four-parameter logistic using GraphPad Prism v.5.

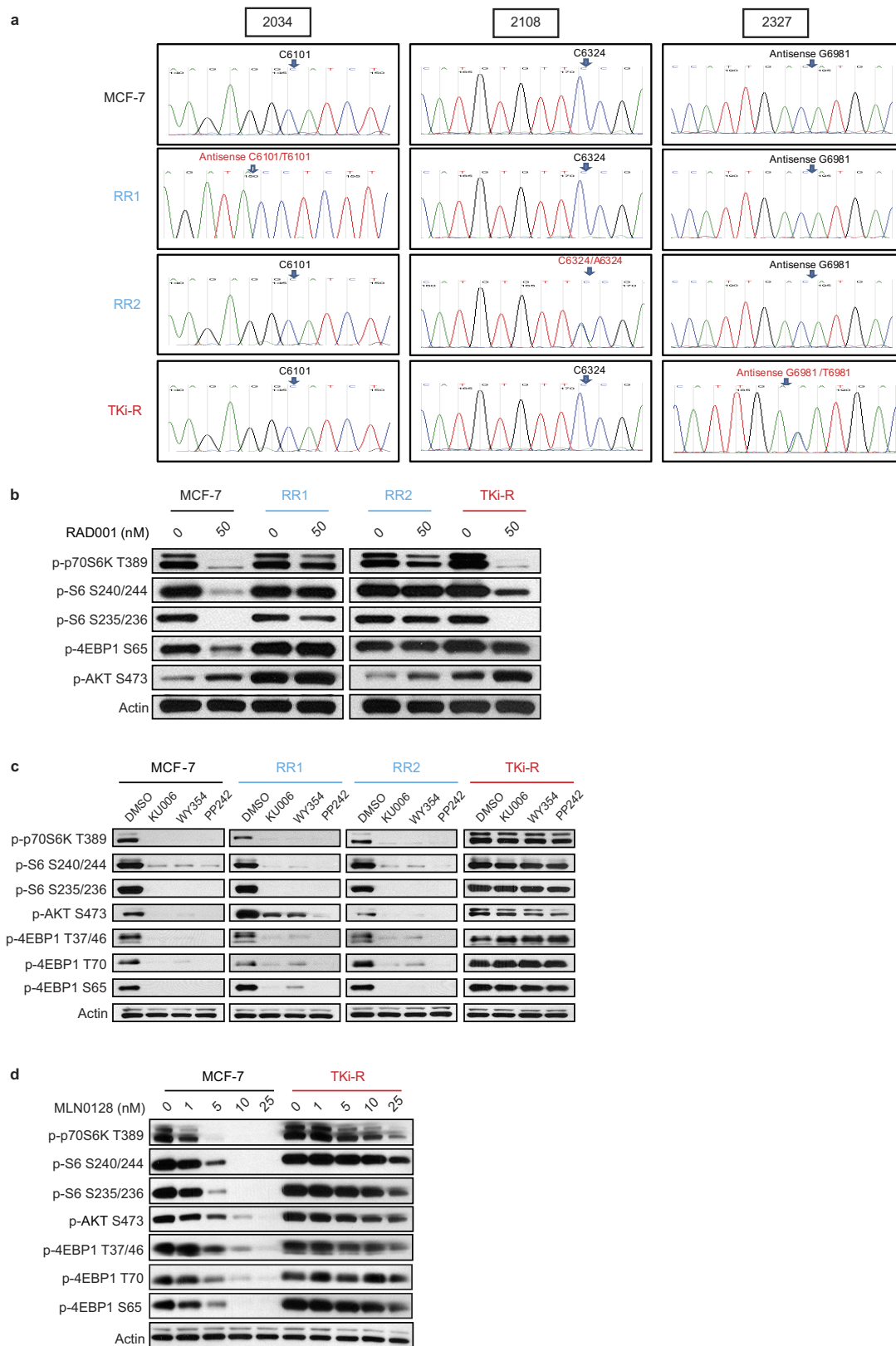
**Animal studies.** All *in vivo* studies were conducted in accordance with guidelines approved by the Memorial Sloan-Kettering Cancer Center (MSKCC) Institutional Animal Care and Use Committee (IACUC). The maximal tumour volume permitted by the MSKCC IACUC is 2,000 mm<sup>3</sup>; this limit was not exceeded in any of the experiments. Eight-week-old athymic nu/nu female mice (Harlan Laboratories) were injected subcutaneously with 10 million cells together with matrigel (BD Biosciences). 17β-Oestradiol pellets (0.72 mg/90 days release) (Innovative Research of America) were implanted subcutaneously 3 days before tumour cell inoculation. Once tumours reached an average volume of 100 mm<sup>3</sup>, mice were randomized (*n* = 5 mice per group) to receive rapamycin (10 mg kg<sup>-1</sup>), AZD8055 (75 mg kg<sup>-1</sup>), RapaLink-1 (1.5 mg kg<sup>-1</sup>) or vehicle only as control. Sample size was chosen based on previous experiments. Rapamycin was formulated in DMSO and delivered



intraperitoneally, AZD8055 was formulated in 30% captisol, and administered orally, RapaLink-1 was formulated in 6% DMSO and 30% captisol and delivered intraperitoneally. Mice treated with RapaLink-1 were also given subcutaneous saline injections twice a day, along with water supplemented with 5% glucose. Tumours were measured twice weekly using callipers, and tumour volume was calculated using the formula:  $\text{length} \times \text{width}^2 \times 0.52$ . Samples were lysed and processed as previously described<sup>29</sup>.

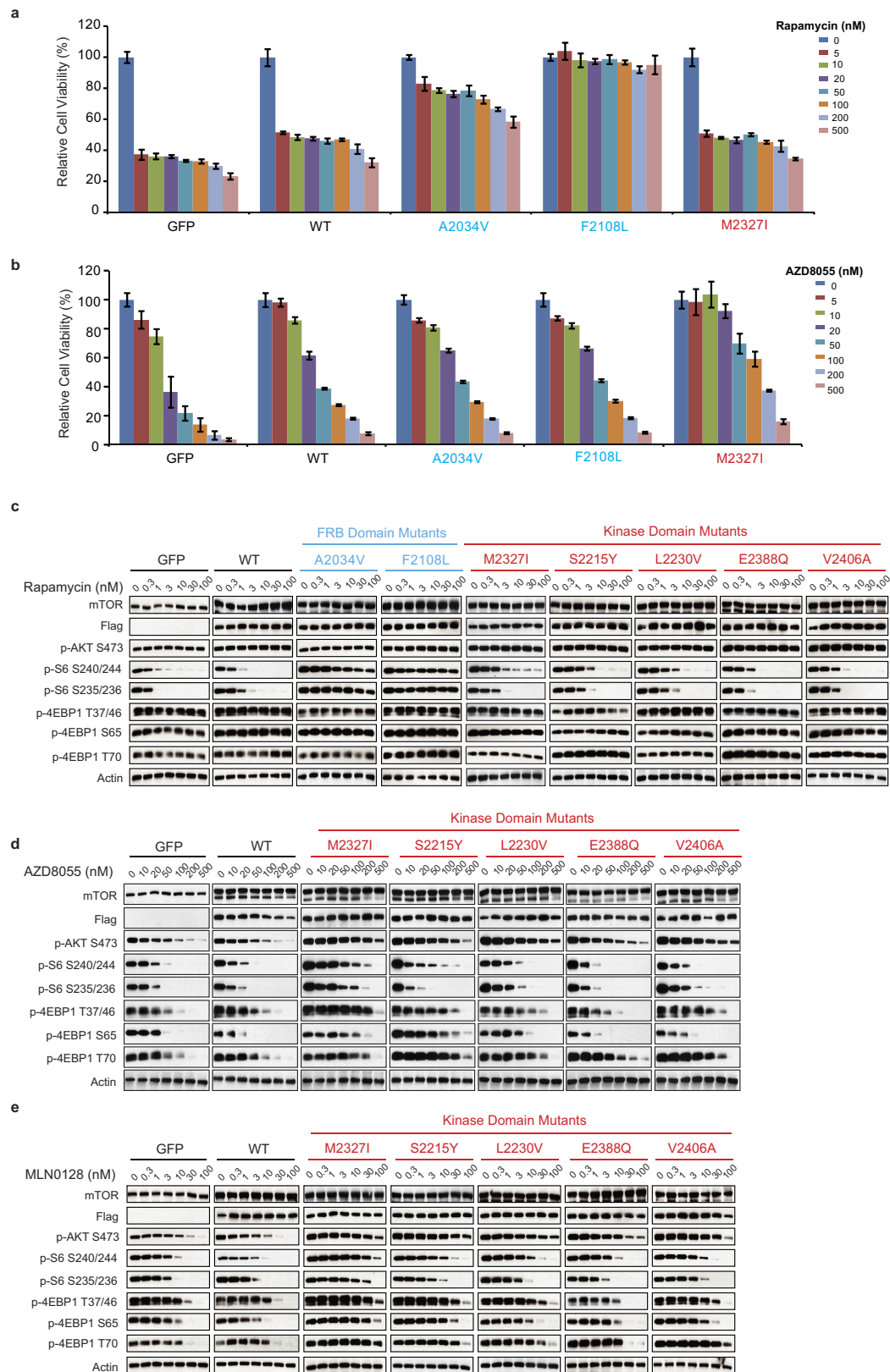
**Statistical analysis.** Results are mean values  $\pm$  s.d. Investigators were not blinded when assessing the outcome of the *in vivo* experiments. All cellular experiments were repeated at least three times.

26. Juric, D. *et al.* Convergent loss of PTEN leads to clinical resistance to a PI(3)K $\alpha$  inhibitor. *Nature* **518**, 240–244 (2015).
27. Yao, Z. *et al.* BRAF mutants evade ERK-dependent feedback by different mechanisms that determine their sensitivity to pharmacologic inhibition. *Cancer Cell* **28**, 370–383 (2015).
28. Cheng, A. C., Eksterowicz, J., Geuns-Meyer, S. & Sun, Y. Analysis of kinase inhibitor selectivity using a thermodynamics-based partition index. *J. Med. Chem.* **53**, 4502–4510 (2010).
29. Rodrik-Outmezguine, V. S. *et al.* mTOR kinase inhibition causes feedback-dependent biphasic regulation of AKT signaling. *Cancer Discov.* **1**, 248–259 (2011).



**Extended Data Figure 1 | Acquired-mTOR mutations promote resistance to mTOR inhibitors in MCF-7 cells. a,** The RNA from MCF-7 parental, RR1, RR2 and TKi-R cells was isolated and the polymerase chain reaction with reverse transcription (RT-PCR) products were submitted to Sanger sequencing at Genewiz. **b,** MCF-7 parental, RR1, RR2 and TKi-R cells were treated with either dimethylsulfoxide (DMSO) or 50 nM

of RAD001 for 4 h. Immunoblot analyses were performed on mTOR effectors. **c, d,** MCF-7 parental, RR1, RR2 and TKi-R cells were treated with either DMSO as a control or 500 nM of either KU006, WY354 or PP242 mTOR inhibitors (**c**), or with different doses of MLN0128 (**d**) for 4 h. Immunoblot analyses were performed on mTOR effectors. All cellular experiments were repeated at least three times.

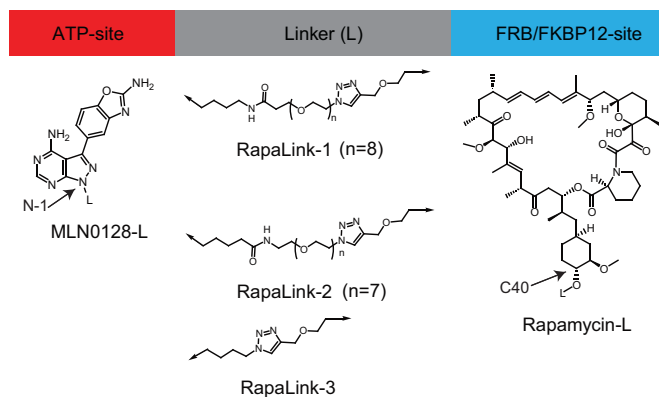


**Extended Data Figure 2 | Acquired-mTOR mutations promote resistance to mTOR inhibitors in MDA-MB-468 cells a, b,** Dose-dependent cell growth inhibition of the MDA-MB-468 cells expressing green fluorescent protein (GFP), wild-type mTOR or different mTOR variants (A2034V, F2108L and M2327I) upon rapamycin (a) or AZD8055 treatment (b). Cells were pre-treated for 24 h with doxycycline ( $1 \mu\text{g ml}^{-1}$ )

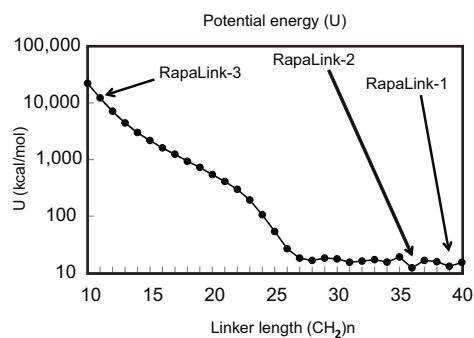
to induce the expression of exogenous mTOR. The cell growth was determined as described in Fig. 1d. c–e, MDA-MB-468 cells expressing GFP, wild-type mTOR or different mTOR variants were treated with different concentrations of rapamycin (c), AZD8055 (d) or MLN0128 (e) for 4 h. Immunoblot analyses were performed on mTOR effectors. All cellular experiments were repeated at least three times.



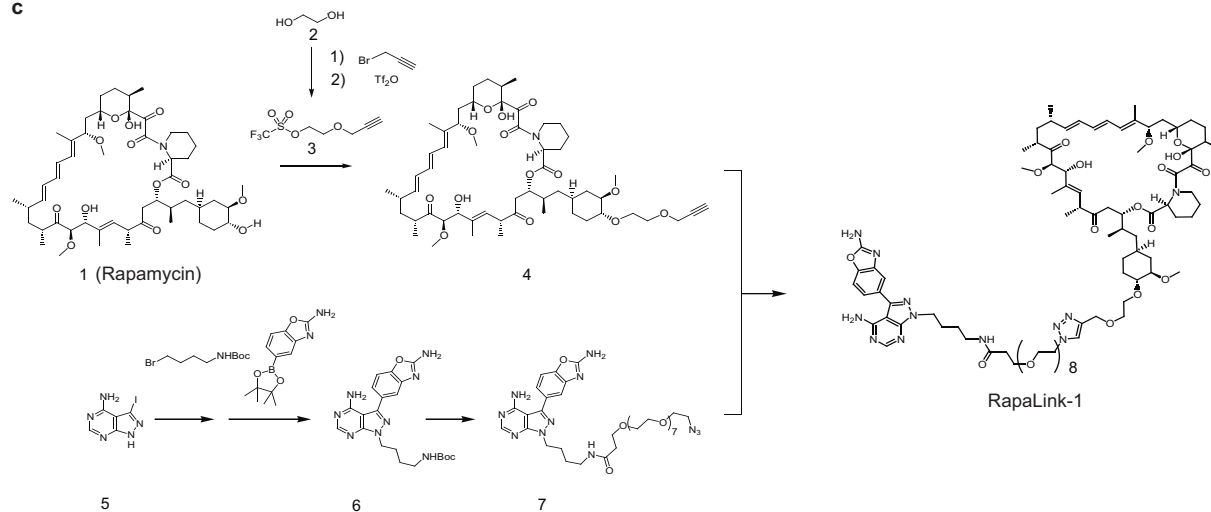
a



b

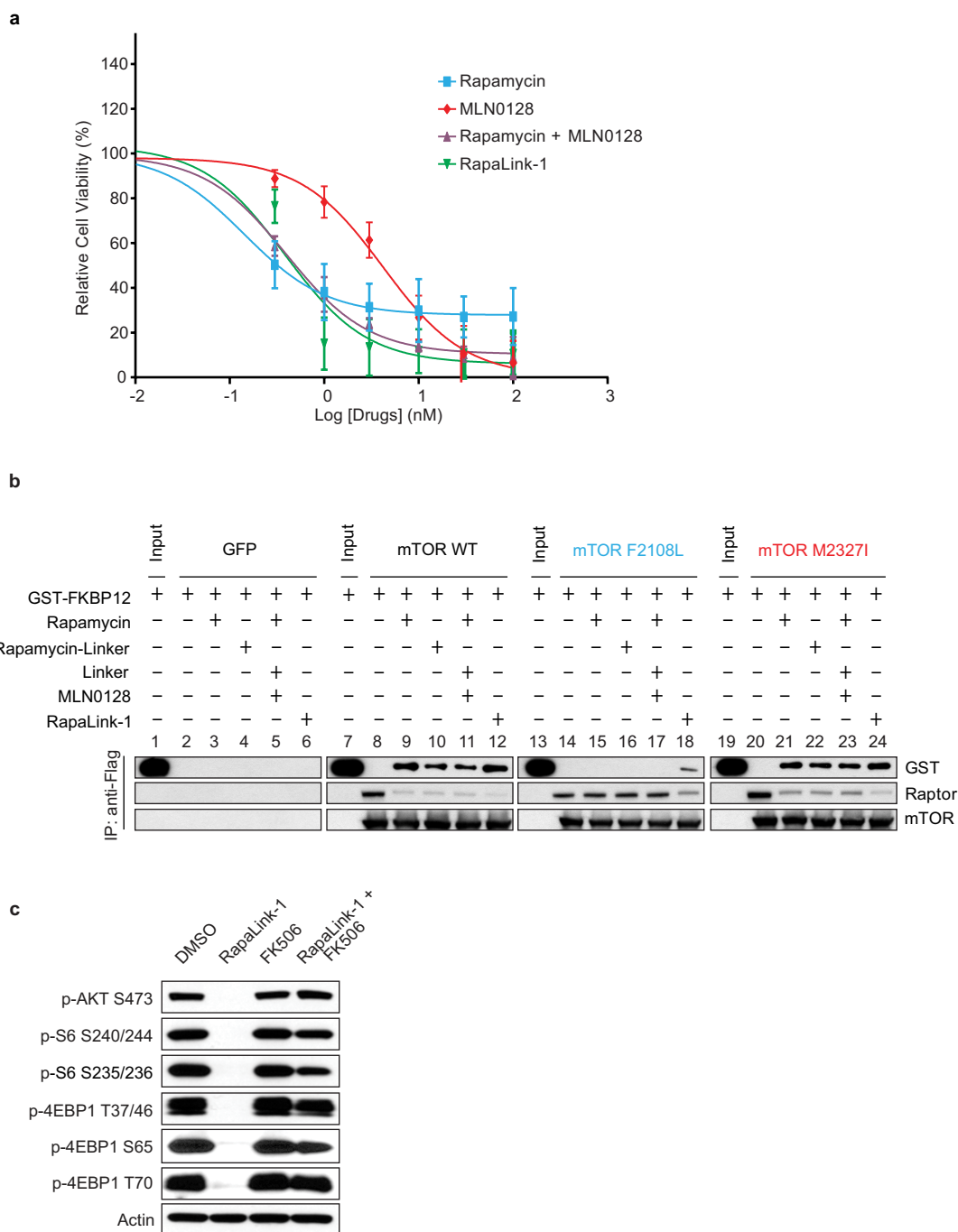


c



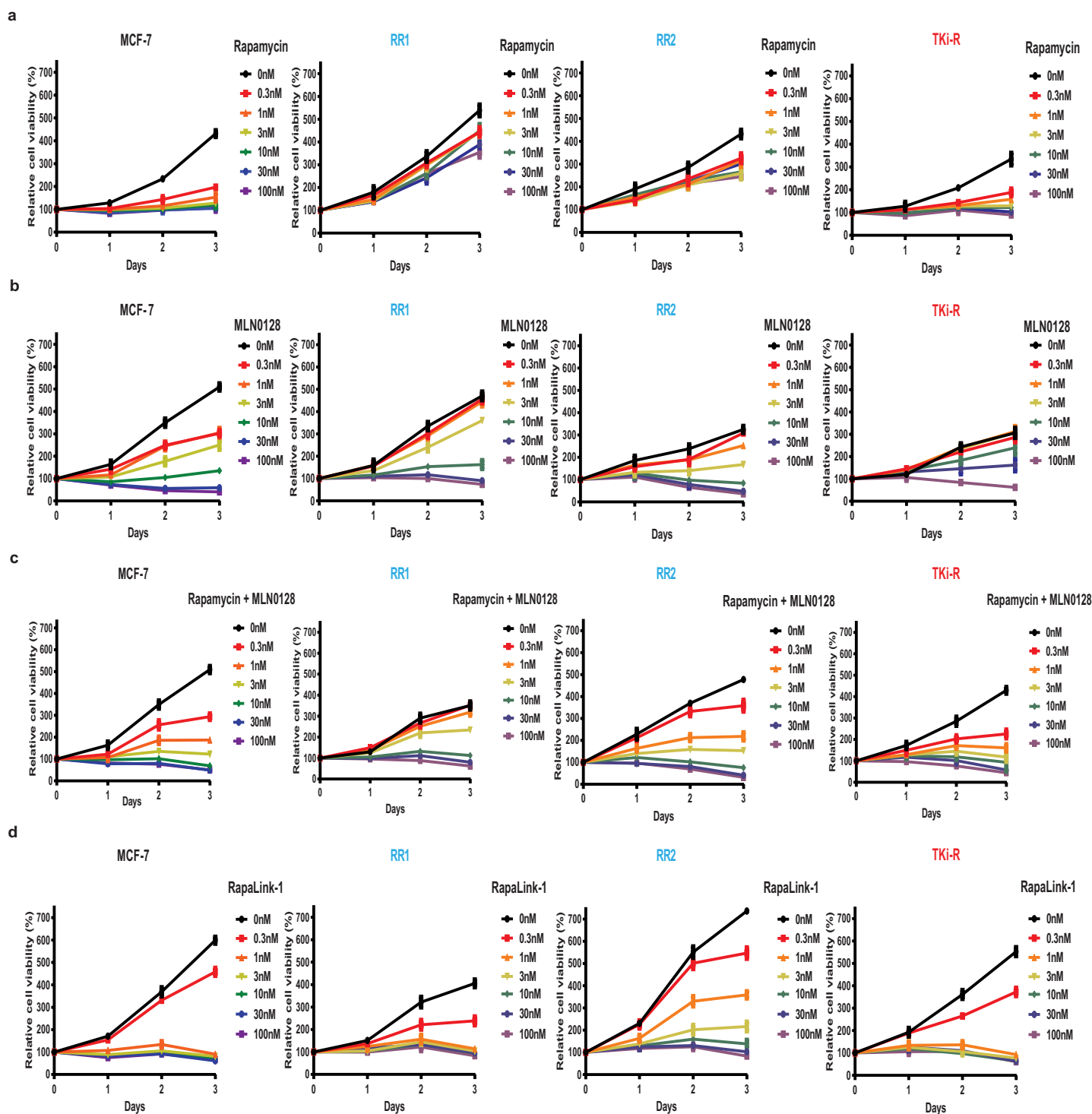
**Extended Data Figure 3 | Synthesis of the mTOR bivalent inhibitor RapaLink-1.** **a**, Compound design of RapaLink-1, -2, and -3 possessing a polyethylene glycol unit of varying lengths. **b**, Calculated potential energy units (U) ( $\text{kcal mol}^{-1}$ ) of modelled compounds of varying methylene

( $\text{CH}_2$ ) $_n$  linker lengths for bivalent interactions with the catalytic site and the FKBP12 site. **c**, A convergent synthetic route for a bivalent mTOR inhibitor RapaLink-1.



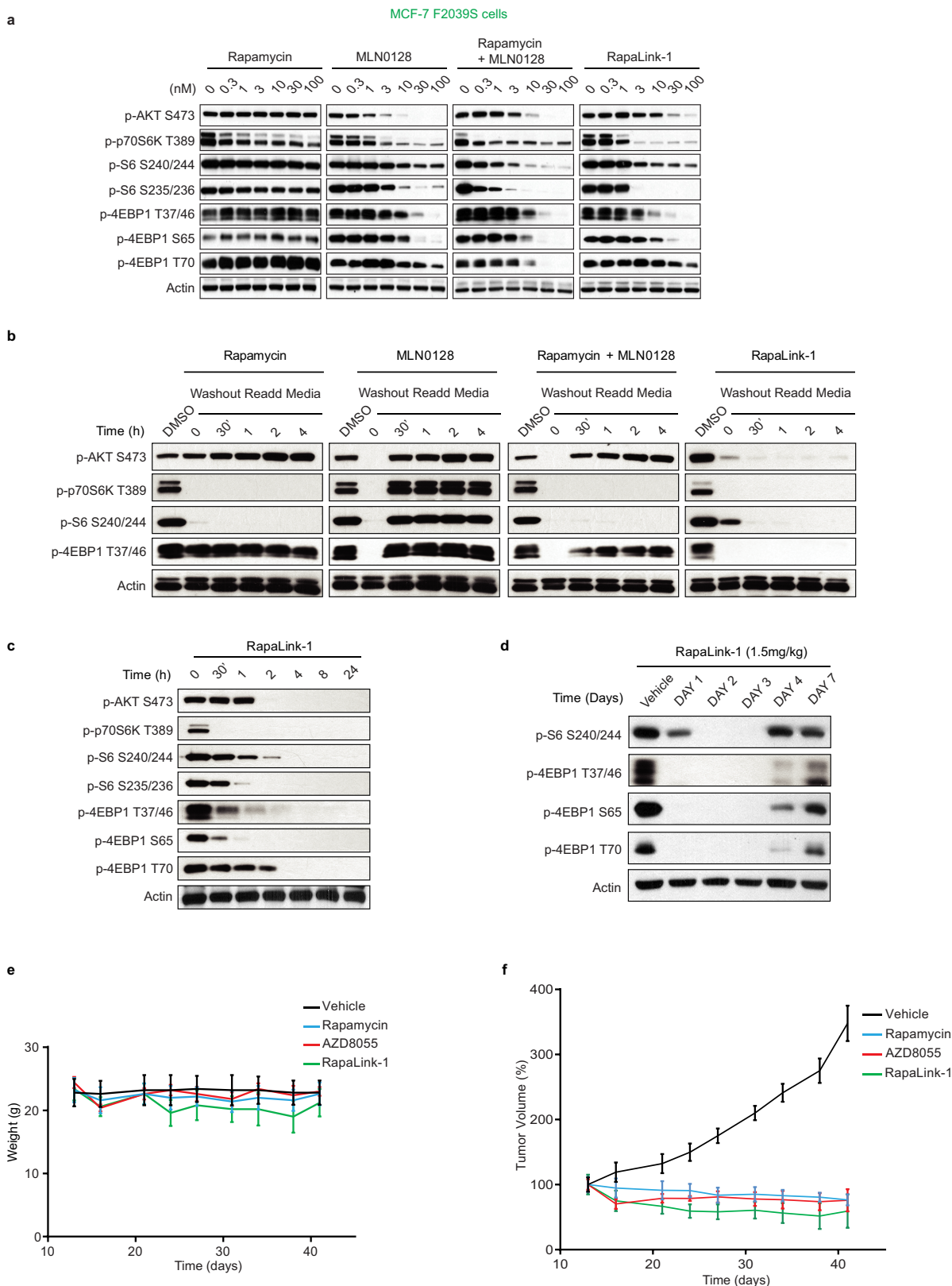
**Extended Data Figure 4 | RapaLink-1 requires FKBP12 for binding to the mTOR FRB domain.** **a**, Dose-dependent cell growth inhibition curves of the MCF-7 parental cell line treated with rapamycin, MLN0128, a combination of rapamycin and MLN0128, or RapaLink-1. The cell growth was determined as described in Fig. 1d. **b**, mTOR-Flag wild type and variants were transfected into 293H cells. The mTORC1 complex was

isolated, and an *in vitro* competition assay in the presence of FKBP12 was performed as described in Fig. 2b. **c**, MCF-7 cells were treated with either DMSO, RapaLink-1 (10 nM), FK506 (10  $\mu$ M), or a combination of both for 24 h, at which time the cells were collected. Immunoblot analyses were performed on mTOR signalling. All experiments were repeated at least three times.



**Extended Data Figure 5 | RapaLink-1 is a potent mTOR inhibitor in wild-type and mutant mTOR cells.** a–d, MCF-7, RR1, RR2 and TKi-R cells were treated with different concentrations of rapamycin (a), MLN0128 (b), combination treatment (c) or RapaLink-1 (d) over 3 days. The cell growth was determined as described in Fig. 1d. Each dot and error bar on the curves represents mean  $\pm$  s.d. ( $n = 8$ ).

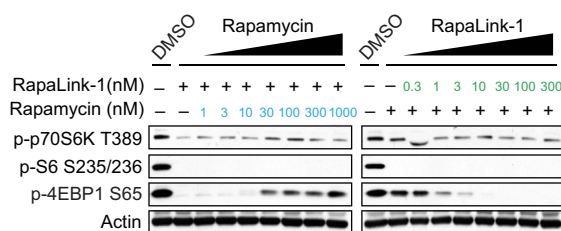




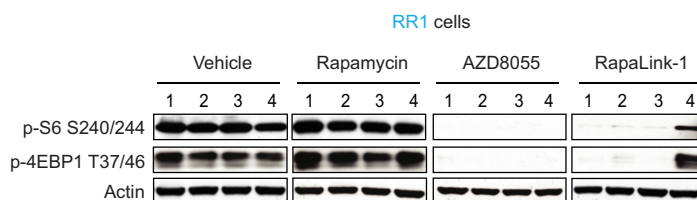
**Extended Data Figure 6 | RapaLink-1 has a prolonged intracellular half-life in wild-type mTOR cells.** **a**, MCF-7 F2039S cells were treated with different concentrations of rapamycin, MLN0128, combination treatment or RapaLink-1 for 4 h, at which time the cells were collected. Immunoblot analyses were performed on mTOR signalling. **b**, MCF-7 cells were treated for 4 h with either DMSO control, 30 nM of rapamycin, 30 nM of MLN0128, a combination of 30 nM of both or 30 nM of RapaLink-1 for 4 h, at which time the treatments were washed out three times with PBS and fresh media was re-added for the indicated times. Immunoblot analyses were performed on mTOR effectors. **c**, MCF-7 cells were

treated with 10 nM of RapaLink-1 and collected at the indicated times. Immunoblot analyses were performed as described earlier. All experiments were repeated at least three times. **d**, Mice bearing MCF-7 xenograft tumours were treated with one single dose of vehicle or RapaLink-1 ( $1.5 \text{ mg kg}^{-1}$ ), tumours were collected at different days after treatment as indicated. Immunoblot analyses were performed on mTOR effectors. **e**, The weight of the mice treated in the efficacy study shown in **f** is reported here. **f**, Mice bearing MCF-7 xenograft tumours were treated as described in Fig. 4c ( $n = 5$  for each group). The results were reported as percentage tumour volume  $\pm$  s.d.

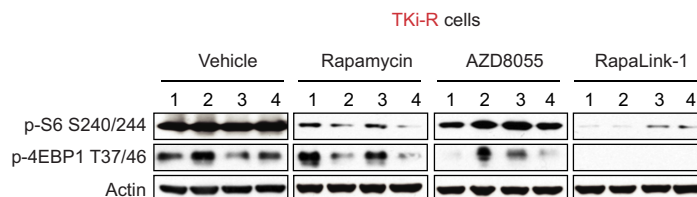
a



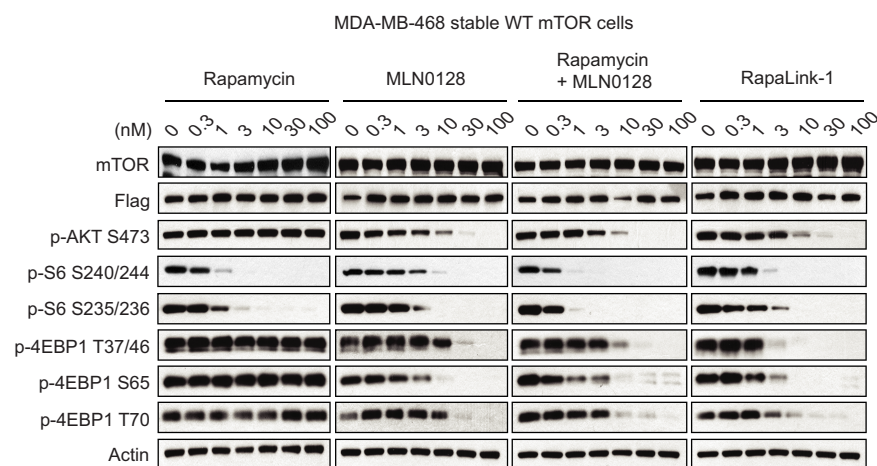
b



c



d



**Extended Data Figure 7 | Rapalink-1 is a more potent mTOR inhibitor than rapamycin.** **a**, MCF-7 cells were treated for 4 h with either Rapalink-1 (10 nM) or rapamycin (10 nM) with simultaneous addition of increasing doses of either rapamycin (left) or Rapalink-1 (right). Immunoblot analyses were performed on mTOR effectors. **b**, **c**, Mice bearing RR1 (**b**) or TKi-R (**c**) xenograft tumours were treated for 24 h with a single dose of either vehicle, rapamycin (10 mg kg<sup>-1</sup>), AZD8055 (75 mg kg<sup>-1</sup>)

or Rapalink-1 (1.5 mg kg<sup>-1</sup>) ( $n = 4$  for each group). Immunoblot analyses were performed on mTOR effectors. **d**, MDA-MB-468 cells inducibly expressing mTOR wild type were treated with either rapamycin, MLN0128, a combination of rapamycin and MLN0128, or Rapalink-1 for 4 h. Immunoblot analyses were performed on mTOR effectors with the indicated antibodies. Rapamycin and MLN0128 panels are the same shown for wild type in Extended Data Fig. 2c and e, respectively.

**Extended Data Table 1 | mTOR mutations found in human patient samples**

References	Histology	mTOR mutations	Domains	Treatments
TCGA-B0-4852-01	ccRCC (TCGA)	E2033V	FRB	Baseline
Wagle et al. 2014	Thyroid	F2108L	FRB	Everolimus
TCGA-DU-6393-01	Glioma (TCGA)	M2327I	Kinase	Baseline
TCGA-A3-3347-01	ccRCC (TCGA)	M2327I	Kinase	Baseline
P-00006559-T01-IM5	Colorectal Cancer (MSKCC-IMPACT)	M2327I	Kinase	Baseline
P-0000645-T01-IM3	Bladder Urothelial Carcinoma (MSKCC-IMPACT)	M2327I	Kinase	Baseline
ccRCC_28	ccRCC (U Tokyo)	M2327I	Kinase	Baseline
P-0000614-T01-IM3	Endometrial Cancer (MSKCC-IMPACT)	M2327V	Kinase	Baseline

List of some FRB and kinase domain mTOR mutations found in human patient samples. Data were collected from the cBioPortal, Memorial Sloan-Kettering Cancer Center (MSKCC).

Extended Data Table 2 | List of FRB domain mutations found in human patient samples

References	Cancer Study	Mutations
P-0004920-T01IM5	MSK-IMPACT	E2014K
SNU738_CENTRAL	CCLE	L2016R
BL41_HAEATOPOIE...	CCLE	I2017T
TCGA-KM-844101	chRCC (TCGA)	I2017T
P-0003073-T01IM5	MSK-IMPACT	I2017S
TCGA-66-2787-01	Lung squ (TCGA)	W2023G
BCP1_HAEMATOPOIE...	CCLE	G2030V
TCGA-B0-4852-01	ccRCC (TCGA)	E2033V
TCGA-B5-A1E-01	Uterine (TCGA)	A2056V
TCGA-G3-A25U-01	Liver (TCGA)	M2057I
TCGA-AZ-6598-01	Colorectal (TCGA)	Q2063fs
H072999	Liver (AMC)	K2066R
NCIH792_LUNG	CCLE	Q2072
P-0000112-T01IM3	MSK-IMPACT	Q2072R
S05-31806-TP-NT	CSCC (Dana-Farber)	X2073_splice
P-0005266-T01IM5	MSK-IMPACT	R2076L
TCGA-EE-A2GC-06	Melanoma (TCGA)	D2077E
P-0002050-T01IM3	MSK-IMPACT	W2084C
TCGA-AX-A0J0	Uterine (TCGA)	Y2088D
TCGA-CQ-5327	Head & neck (TCGA)	M2089I
P-0005159-T01IM5	MSK-IMPACT	L2097F
TCGA-D8-A1Y0-01	Breast (TCGA)	W2101L

Data were collected from the cBioPortal, MSKCC.



Extended Data Table 3 | List of mTOR kinase domain mutations found in human patient samples

References	Cancer Study	Mutations
TCGA-C5-A1BM-01	Cervical (TCGA)	D2191H
TCGA-BR-4370-01	Stomach (TCGA)	R2193C
P-0004376-T01-IM5	MSK-IMPACT	F2202L
TCGA-BT-A0YX-01	Bladder (TCGA)	L2209V
TCGA-B0-4810-01	ccRCC (TCGA)	A2210P
JHUEM7_ENDOMETRIUM	CCLE	S2215Y
TCGA-JW-A5VL-01	Cervical (TCGA)	S2215Y
TCGA-FU-A3HZ-01	Cervical (TCGA)	S2215Y
TCGA-AA-A00K-01	Colorectal (TCGA)	S2215Y
TCGA-A6-6141-01	Colorectal (TCGA)	S2215Y
TCGA-F4-6806-01	Colorectal (TCGA)	S2215F
TCGA-DM-A1D4	Colorectal (TCGA)	S2215F
S12-23181-TP-NT	CSCC (Dana-Farber)	S2215F
TCGA-CJ-5679-01	ccRCC (TCGA)	S2215Y
TCGA-A4-7828-01	pRCC (TCGA)	S2215Y
P-0000208-T01-IM3	MSK-IMPACT	S2215Y
P-0000208-T02-IM5	MSK-IMPACT	S2215Y
P-0005214-T01-IM5	MSK-IMPACT	S2215Y
MEL-UKRV-Mel-20	Melanoma (Broad)	S2215Y
TCGA-BS-A0UF-01	Uterine (TCGA)	S2215Y
TCGA-BS-A0UV-01	Uterine (TCGA)	S2215Y
TCGA-BG-A0VX	Uterine (TCGA)	S2215Y
TCGA-CS-5396-01	Glioma (TCGA)	L2216P
HEC251_ENDOMETRIUM	CCLE	R2217V
TCGA-A6-4105-01	Colorectal (TCGA)	Q2223K
TCGA-CJ-4887-01	ccRCC(TCGA)	L2230V
SNU1196_BILIARY	CCLE	S2231W
KMH2_HAEMATOPHIE	CCLE	T2232I
CW2_LARGE_INTEST	CCLE	G2238D
TCGA-JW-A5VL-01	Cervical (TCGA)	W2239C
P-0000614-T01-IM3	MSK-IMPACT	W2239Ffs*39
DS-bla-084	Bladder (MSKCC 2014)	P2241S
MSKCC-0296_R	Bladder (DFARBER MSKCC 2014)	P2241S
TCGA-DK-A6B6-01	Bladder (TCGA)	P2241S
TCGA-JW-A5VL-01	Cervical (TCGA)	P2241S
TCGA-97-7554-01	Lung adeno (TCGA)	R2251Q
HEC1A_ENDOMETRIUM	CCLE	R2254M
SNU520_STOMACH	CCLE	H2265P
TCGA-75-5126-01	Lung adeno (TCGA)	R2266P
TCGA-34-5239-01	Lung squ (TCGA)	A2272S
TCGA-EB-A3Y7-01	Melanoma (TCGA)	A2272V
P-0005309-T01-IM5	MSK-IMPACT	Q2282P
TCGA-BR-8363-01	Stomach (TCGA)	T2294A
P-0001171-T01-IM3	MSK-IMPACT	D2298H
DS-bla-037	Bladder (MSKCC 2014)	E2311K
P-0001042-T01-IM3	MSK-IMPACT	A2325V
P-0003529-T01-IM5	MSK-IMPACT	V2330I
TCGA-AA-3666-01	Colorectal (TCGA)	I2333M
TCGA-B0-5691-01	Renal clear cell (TCGA)	L2334V
TCGA-HU-A4GU-01	Stomach (TCGA)	R2339_splice
L540_HAEMATOPHIE	CCLE	H2340R
TCGA-HB-A43Z-01	Sarcoma (TCGA)	S2342Y
NCIH446_LUNG	CCLE	M2345V
ESCC-148T	Esophagus sq (ICGC)	G2351E
LUAD-FH5PJ	Lung adeno (BROAD)	H2355R
S12-23181-TP-NT	CSCC (Dana-Farber)	G2359R
TCGA-E9-A54Y-01	Breast (TCGA)	R2368Q
HEC108-ENDOMETRIUM	CCLE (Novartis/Broad 2012)	R2368Q
P-0002079-T01-IM3	MSK-IMPACT	L2383F
P-0001042-T01-IM3	MSK-IMPACT	T2384I
TCGA-BH-A0HP-01	Breast (TCGA)	E2388Q
SW48_LARGE_INTES	CCLE	V2389_splice
COSM51966	Kidney	V2406A
TCGA-A7-A5ZV-01	Breast (TCGA)	D2412H
H061142	Liver (AMC)	D2412V
TCGA-EK-A3GK-01	Cervical (TCGA)	E2419K
CLL147	CLL (BROAD)	A2420P
P-0003254-T01-IM5	MSK-IMPACT	A2420V
P-0004798-T01-IM5	MSK-IMPACT	A2420V
TCGA-FS-A4FC-06	Melanoma (TCGA)	A2420V
TCGA-ER-A42L-06	Melanoma (TCGA)	D2424N
ST486-HAEMATOPHIE	CCLE	L2427R
TCGA-06-0122-01	GBM (TCGA)	L2427Q
TCGA-KN-8437-01	chRCC (TCGA)	L2427R
P-0005355-T01-IM5	MSK-IMPACT	L2427R
P-0005044-T01-IM5	MSK-IMPACT	L2427Q
P-0000839-T01-IM3	MSK-IMPACT	L2427Q
MOLT16_HAEMATOP	CCLE	R2430M
TCGA-G4-6322-01	Colorectal (TCGA)	R2441Q

Data were collected from the cBioPortal, MSKCC.

# Ribosome-dependent activation of stringent control

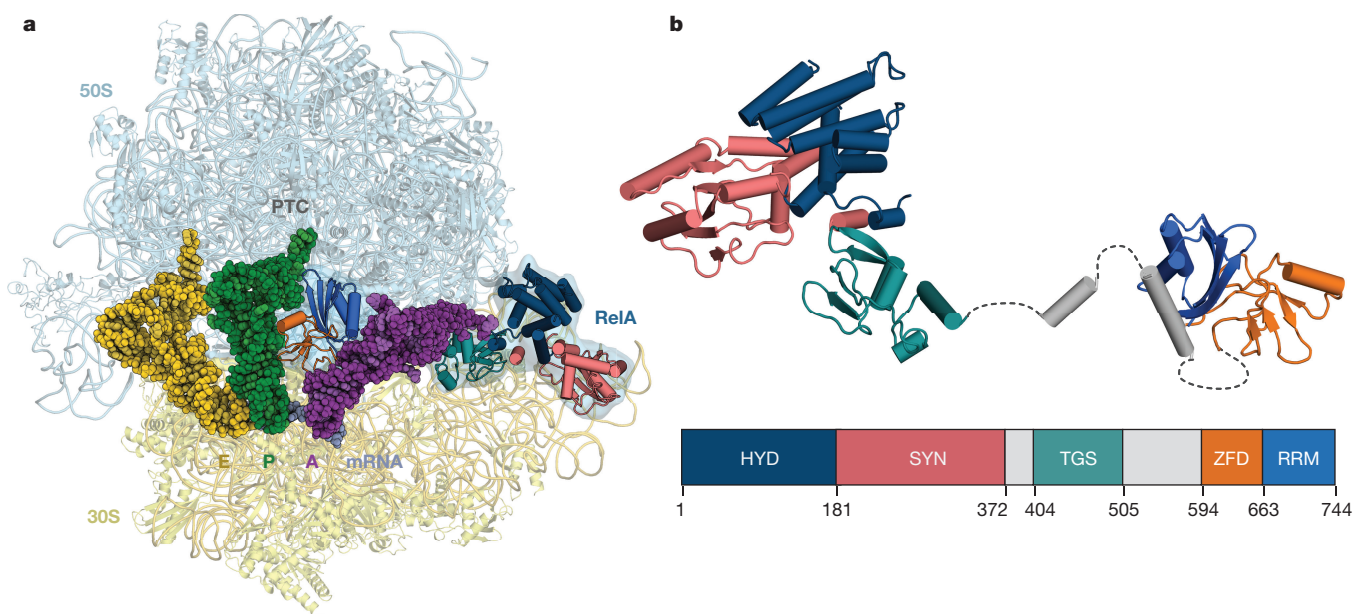
Alan Brown<sup>1\*</sup>, Israel S. Fernández<sup>1†\*</sup>, Yuliya Gordiyenko<sup>1</sup> & V. Ramakrishnan<sup>1</sup>

In order to survive, bacteria continually sense, and respond to, environmental fluctuations. Stringent control represents a key bacterial stress response to nutrient starvation<sup>1,2</sup> that leads to rapid and comprehensive reprogramming of metabolic and transcriptional patterns<sup>3</sup>. In general, transcription of genes for growth and proliferation is downregulated, while those important for survival and virulence are upregulated<sup>4</sup>. Amino acid starvation is sensed by depletion of the aminoacylated tRNA pools<sup>5</sup>, and this results in accumulation of ribosomes stalled with non-aminoacylated (uncharged) tRNA in the ribosomal A site<sup>6,7</sup>. RelA is recruited to stalled ribosomes and activated to synthesize a hyperphosphorylated guanosine analogue, (p)ppGpp<sup>8</sup>, which acts as a pleiotropic secondary messenger. However, structural information about how RelA recognizes stalled ribosomes and discriminates against aminoacylated tRNAs is missing. Here we present the cryo-electron microscopy structure of RelA bound to the bacterial ribosome stalled with uncharged tRNA. The structure reveals that RelA utilizes a distinct binding site compared to the translational factors, with a multi-domain architecture that wraps around a highly distorted A-site tRNA. The TGS (ThrRS, GTPase and SpoT) domain of RelA binds the CCA tail to orient the free 3' hydroxyl group of the terminal adenosine towards a  $\beta$ -strand, such that an aminoacylated tRNA at this position would be sterically precluded. The structure

supports a model in which association of RelA with the ribosome suppresses auto-inhibition to activate synthesis of (p)ppGpp and initiate the stringent response. Since stringent control is responsible for the survival of pathogenic bacteria under stress conditions, and contributes to chronic infections and antibiotic tolerance, RelA represents a good target for the development of novel antibacterial therapeutics.

Stringent control is a pleiotropic response to the failure of amino acid availability to keep up with the demands of protein synthesis<sup>1</sup>. It is mediated by a hyperphosphorylated nucleotide ((p)ppGpp)<sup>9,10</sup>. In *E. coli*, the synthesis of (p)ppGpp is catalysed by RelA<sup>2</sup>, a multi-domain ATP:GTP(GDP) pyrophosphate transferase, and a prototypic member of the RelA/SpoT homologue (RSH) family<sup>11</sup>. The majority of multi-domain RSH proteins are stimulated to generate (p)ppGpp in a ribosome-dependent manner when an uncharged and cognate tRNA, which acts as a marker for nutrient deficiency, is located in the ribosomal A site<sup>7,8</sup>. Discrimination against aminoacylated tRNA prevents undesired activation of stringent control during the normal translation cycle.

Using cryo-electron microscopy we have solved the structure of the *E. coli* ribosome, programmed so that uncharged tRNA(Phe) occupies the A site, in complex with RelA at an overall resolution of 3.0 Å (Fig. 1, Extended Data Figs 1 and 2, and Extended Data Table 1). We did not



**Figure 1 | Structure of RelA bound to the ribosome.** **a**, Overall view of RelA in complex with a ribosome stalled with an uncharged tRNA in the A site. Displayed are the 50S and 30S ribosomal subunits; E-, P- and A-site tRNAs; mRNA, and RelA coloured by domain. **b**, Structure of the ribosome-bound form of RelA oriented from N to C termini with the

domain organization below showing the boundaries of the hydrolase (HYD), synthetase (SYN), TGS, zinc-finger (ZFD) and RNA recognition motif (RRM) domains. Unmodelled flexible elements that connect RelA domains are indicated with dashed lines.

<sup>1</sup>MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, UK.

<sup>†</sup>Present address: Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032, USA.

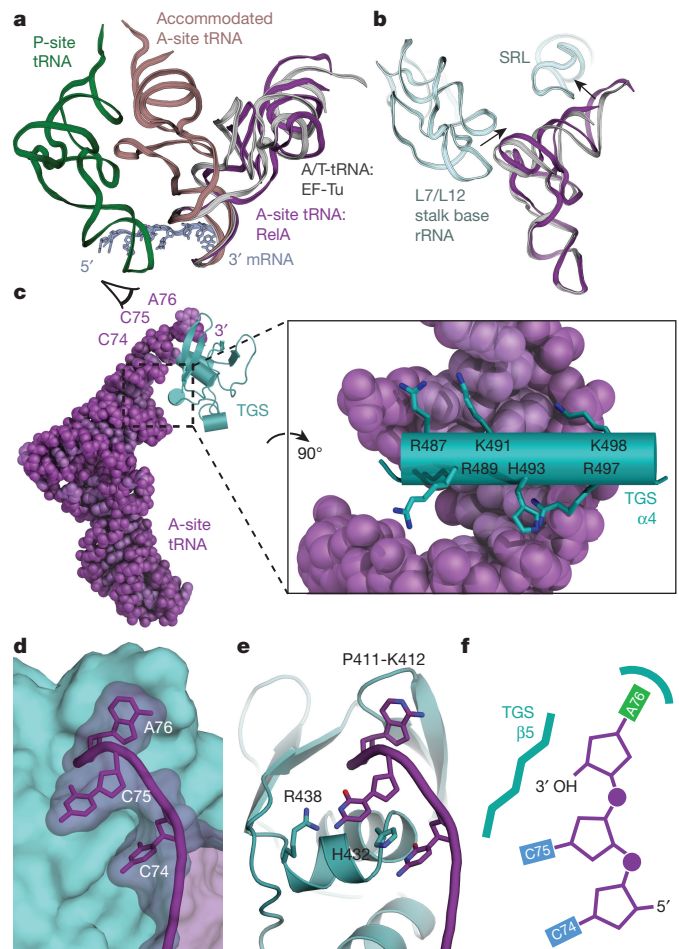
\*These authors contributed equally to this work

observe any class in which RelA was bound to the ribosome in the absence of A-site tRNA. Both RelA and the A-site tRNA remain flexible when bound to the ribosome, primarily due to binding intrinsically flexible rRNA elements (notably the L7/L12 stalk base and the A-site finger). Although there are only minor differences in conformations (Extended Data Fig. 1b), the heterogeneity was sufficient to result in RelA having less well-resolved density than the ribosome. To distinguish conformational states and improve the local map quality we used a recent modification of the 3D classification process<sup>12</sup>, in which ribosome projections were subtracted from each experimental particle, leaving signal only for RelA before classification focused on each domain (Methods and Extended Data Fig. 1). This improved the density for the RelA domains (Extended Data Figs 2 and 3) allowing models to be built (Extended Data Table 2).

The structure reveals that RelA forms a highly extended conformation on the ribosome to cradle the uncharged tRNA in a distorted conformation in the A site (Fig. 1). RelA has an N-terminal region formed by hydrolase (HYD), synthetase (SYN), and TGS domains that are located at the acceptor end of the A-site tRNA, and a C-terminal region formed by a zinc-finger domain and an RNA recognition motif (RRM) that run parallel to the anticodon arm of the tRNA. These five domains are connected by flexible and helical elements in a serpentine configuration that wind between the ribosome and the A-site tRNA (Fig. 1 and Extended Data Fig. 4). In this conformation, RelA inhibits accommodation of the acceptor arm of the uncharged tRNA into the peptidyl transferase centre.

As has been previously noted<sup>13</sup>, the overall conformation of the A-site tRNA in the presence of RelA resembles the A/T state adopted by pre-accommodated aminoacyl-tRNA in complex with EF-Tu<sup>14</sup> (Fig. 2a). However, our high-resolution map reveals that the interactions between the tRNA and the ribosomal large subunit are very different, with the tRNA contacting both the RNA component of the L7/L12 stalk base (helices 42–44) and the sarcin-ricin loop (SRL; helix 95) (Fig. 2b). A stacking interaction between nucleotides C56 of the tRNA elbow and A1067 of the L7/L12 stalk base is reminiscent of how the L1 stalk (helices 76–78) recognizes E-site tRNA<sup>15</sup>. The antibiotic thiostrepton binds in the vicinity of A1067 and may prevent this interaction, explaining its ability to inhibit (p)ppGpp formation<sup>8</sup>. The contacts with rRNA, and also with RelA, distort the tRNA compared to the A/T state (Fig. 2a, b and Extended Data Fig. 5). Starting at base-pair 27:43 after the aligned anticodon stem-loops (ASLs), a 6° rotation away from the ribosome is coordinated with an outward movement of the L7/L12 stalk base. A further 11° rotation of the acceptor stem, starting at base-pair 7:66, allows the tRNA to contact the SRL, which is bound by EF-Tu in the decoding complex<sup>14</sup>.

Once EF-Tu has dissociated from the ribosome, it is not known whether the A-site tRNA conformation could still fluctuate or the higher affinity of aminoacylated (compared to uncharged) tRNA for the A site in the peptidyl transferase centre<sup>16,17</sup> would stabilize the accommodated form of tRNA. However, with an uncharged tRNA, the fluctuations in its conformation could bring the acceptor end into contact with the RelA TGS domain and stabilize a distorted form in a manner that discriminates against aminoacylated tRNAs (Fig. 2c). This small domain has a  $\beta$ -grasp fold similar to that found in the ubiquitin family. In RelA, an additional pair of C-terminal  $\alpha$ -helices, the second of which spans across the axis of the acceptor stem, extend the fold (TGS  $\alpha$ 3 and  $\alpha$ 4, see Extended Data Fig. 6). Highly conserved basic residues (Arg487, Lys491, His493, Arg497, and Lys498) at each end of this helix form electrostatic interactions with the tRNA phosphate backbone (Fig. 2c and Extended Data Fig. 7). The absence of base-specific contacts allows RelA to recognize all tRNAs equally. The 3' CCA of the A-site tRNA extends ~14 Å around the outside of the TGS domain (Fig. 2d) and is maintained by a series of interactions; C74 stacks with His432, C75 can form hydrogen bonds with Arg438, and crucially A76 stacks beneath Pro411 and Lys412 (Fig. 2e and Extended Data Fig. 3a–c). This positions the free 3' hydroxyl group



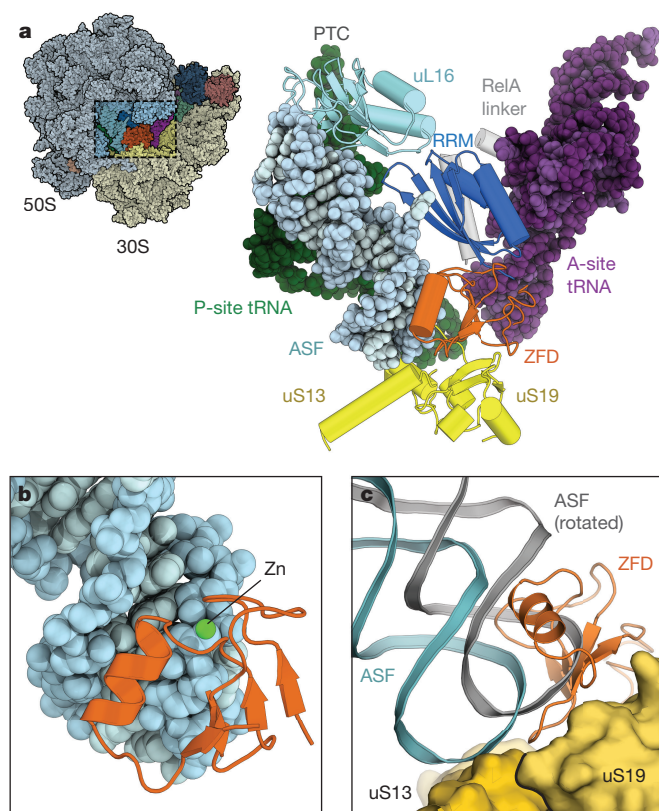
**Figure 2 | Molecular basis for the recognition of uncharged A-site tRNA.** **a**, Comparison of A-site tRNAs. In the presence of RelA, A-site tRNA (purple) structurally resembles the A/T state adopted by pre-accommodated aminoacyl-tRNA in complex with EF-Tu (light grey). RelA prevents accommodation of tRNA into the canonical A-site position (brown). P-site tRNA (green) and mRNA (dark grey) are shown for reference. **b**, Conformational differences between RelA-bound A-site tRNA (purple) and A/T-tRNA (light grey) result from an outward position of the L7/L12 stalk base and movement of the tRNA to contact the SRL (both light blue). **c**, The TGS domain (teal) binds the acceptor end of the A-site tRNA (purple) through a positively-charged helix that interacts with the phosphate backbone (expanded view). **d**, When viewed from the back as indicated in **c**, the 3' CCA (nucleotides 74–76) of the A-site tRNA wraps around the surface of the TGS domain. **e**, The conformation of the CCA is maintained by interactions with invariant residues of the TGS domain. **f**, The free 3' OH of the terminal adenine is positioned to face the  $\beta$ 5-strand of the TGS domain to sterically preclude the binding of aminoacylated tRNAs.

of the terminal adenosine towards the  $\beta$ 5-strand of the TGS domain (Fig. 2e, f). An aminoacylated tRNA at this position would be sterically precluded (Extended Data Fig. 3b). This agrees with data that shows a free 3' hydroxyl group is a prerequisite of RelA activation<sup>18</sup>.

As well as binding the A-site tRNA, the TGS domain contacts the small subunit rRNA and the ribosomal proteins uS12 and uL14. Although ribosomes that lack, or have mutant, uL11 have impaired (p)ppGpp synthesis<sup>19,20</sup>, we do not observe a direct interaction between RelA and uL11.

The TGS domain together with the N-terminal hydrolase and synthetase domains, defines the minimal unit found in ribosome-dependent RSHs<sup>11</sup>. In RelA, the two C-terminal domains act in tandem to further anchor RelA to the ribosome by binding the ribosomal A-site finger (ASF; 23S rRNA helix 38) that spans the inter-subunit space





**Figure 3 | Interactions between RelA and the ribosome.** **a**, Overview (left) and details (right) of the interaction between the zinc-finger domain (ZFD) (orange) and RRM (blue) of RelA and the ribosomal ASF (light blue) that spans the inter-subunit interface between the P-site (green) and A-site (purple) tRNAs. RelA acts as an additional inter-subunit bridge by binding uL16 (cyan) in the large subunit and uS19 (yellow) in the small subunit. PTC, peptidyl transferase centre. **b**, The  $\alpha$ -helix of the ZFD binds in the major groove of the ASF, with the zinc-binding site interacting with the phosphate backbone. **c**, By binding to uS19, the ZFD occupies the position adopted by the ASF in the rotated ribosome (grey).

(Fig. 3a and Extended Data Fig. 7). Although the C terminus was known to be involved in binding to the ribosome<sup>21,22</sup>, the nature of the interaction is unexpected and was not previously noted in the low-resolution reconstruction<sup>13</sup>. To our knowledge, this represents a novel binding site on the ribosome for an extrinsic factor.

RelA interacts with the tip of the ASF through an unpredicted, and unusual, CCHC-type zinc-finger domain. This domain has a  $\beta\beta\beta\alpha\alpha\beta$  topology (Extended Data Fig. 6), with a single zinc ion co-ordinated at the N terminus of the  $\alpha$ 1-helix (Fig. 3b). The zinc-binding residues, Cys612 and His634, directly contribute to binding the ASF by interacting with the phosphate backbone of nucleotide U884 and help to orient the  $\alpha$ 1-helix in the major groove of the ASF. The ASF usually forms a dynamic bridge (B1a) with either uS13 or uS19 depending on the ratcheted state of the ribosome<sup>23</sup>. RelA stabilizes the bridge with uS13 by binding to the same interface of uS19 as occupied by the ASF in the rotated ribosome, which results in a better-defined ASF than in any previous ribosome structure (Fig. 3c).

The C-terminal domain is sandwiched between the upper part of the ASF and the A-site tRNA, and contacts the large subunit protein, uL16 (Fig. 3a). Although this domain had been predicted as an ACT fold<sup>11</sup>, the interaction between the face of its four-stranded antiparallel  $\beta$ -sheet and the ASF is more reminiscent of an RNA recognition motif (RRM), which shares the same topology (Extended Data Fig. 8). Furthermore, conjoint RRM/zinc-finger domains are common in eukaryotes<sup>24</sup>.

In contrast to the C-terminal domains, the hydrolase and catalytic synthetase domains form few contacts with the ribosome. The domains

are extremely flexible, and occupy multiple positions between the SRL and the spur of the small subunit (18S rRNA helix 6). Using our focused classification approach we isolated a few well-populated states that were sufficiently resolved to orient a model based on a crystal structure of a homologue<sup>25</sup> (Extended Data Fig. 1a). Although the synthetase domain can, in some classes, contact the tip of the spur (nucleotides 84–88), the general absence of interactions with the ribosome, suggests that RelA activation is indirect. Rather, as the isolated catalytic domain of RelA is constitutively active in a ribosome-independent manner<sup>22</sup>, our structure supports a model where association with the ribosome and uncharged A-site tRNA suppresses regulatory auto-inhibition that results from RelA homodimers or oligomers<sup>21,26</sup>, as previously suggested<sup>27</sup>. This contrasts with activation of the translational GTPases in which the SRL has a direct role in inducing catalysis<sup>28</sup>. Our structure should provide a framework for the design of experiments to differentiate among the various proposed mechanisms for RelA<sup>20,27,29</sup>.

In conclusion, our structural data reveal how RelA specifically recognizes ribosomes stalled under conditions of amino acid starvation to activate synthesis of (p)ppGpp and initiate the stringent response. By using the ribosome as a signalling platform, RelA provides an immediate link between the status of translation and global adaptation to the environment. As the distribution of the RSH family is strictly limited to bacteria, and stringent control contributes to the virulence, persistence, and antibiotic tolerance of bacterial infections, the structure can provide a framework for the development of therapeutics that can selectively inactivate stringent control and re-sensitize resilient bacteria to antibiotics<sup>30</sup>.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 9 December 2015; accepted 18 March 2016.**

**Published online 9 May 2016.**

- Sands, M. K. & Roberts, R. B. The effects of a tryptophan-histidine deficiency in a mutant of *Escherichia coli*. *J. Bacteriol.* **63**, 505–511 (1952).
- Stent, G. S. & Brenner, S. A genetic locus for the regulation of ribonucleic acid synthesis. *Proc. Natl Acad. Sci. USA* **47**, 2005–2014 (1961).
- Haurlyuk, V., Atkinson, G. C., Murakami, K. S., Tenson, T. & Gerdes, K. Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nat. Rev. Microbiol.* **13**, 298–309 (2015).
- Magnusson, L. U., Farewell, A. & Nyström, T. ppGpp: a global regulator in *Escherichia coli*. *Trends Microbiol.* **13**, 236–242 (2005).
- Fangman, W. L. & Neidhardt, F. C. Protein and ribonucleic acid synthesis in a mutant of *Escherichia coli* with an altered aminoacyl ribonucleic acid synthetase. *J. Biol. Chem.* **239**, 1844–1847 (1964).
- Haseltine, W. A. & Block, R. Synthesis of guanosine tetra- and pentaphosphate requires the presence of a codon-specific, uncharged transfer ribonucleic acid in the acceptor site of ribosomes. *Proc. Natl Acad. Sci. USA* **70**, 1564–1568 (1973).
- Pedersen, F. S., Lund, E. & Kjeldgaard, N. O. Codon specific, tRNA dependent *in vitro* synthesis of ppGpp and pppGpp. *Nat. New Biol.* **243**, 13–15 (1973).
- Haseltine, W. A., Block, R., Gilbert, W. & Weber, K. MSI and MSII made on ribosome in idling step of protein synthesis. *Nature* **238**, 381–384 (1972).
- Cashel, M. & Gallant, J. Two compounds implicated in the function of the RC gene of *Escherichia coli*. *Nature* **221**, 838–841 (1969).
- Cashel, M. & Kalbacher, B. The control of ribonucleic acid synthesis in *Escherichia coli*. V. Characterization of a nucleotide associated with the stringent response. *J. Biol. Chem.* **245**, 2309–2318 (1970).
- Atkinson, G. C., Tenson, T. & Haurlyuk, V. The RelA/SpoT homolog (RSH) superfamily: distribution and functional evolution of ppGpp synthetases and hydrolases across the tree of life. *PLoS One* **6**, e23479 (2011).
- Bai, X.-C., Rajendra, E., Yang, G., Shi, Y. & Scheres, S. H. Sampling the conformational space of the catalytic subunit of human  $\gamma$ -secretase. *elife* **4**, e11182 (2015).
- Agirrezabala, X. *et al.* The ribosome triggers the stringent response by RelA via a highly distorted tRNA. *EMBO Rep.* **14**, 811–816 (2013).
- Schmeing, T. M. *et al.* The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science* **326**, 688–694 (2009).
- Selmer, M. Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313**, 1935–1942 (2006).
- Lill, R., Robertson, J. M. & Wintermeyer, W. Affinities of tRNA binding sites of ribosomes from *Escherichia coli*. *Biochemistry* **25**, 3245–3255 (1986).
- Schilling-Bartetzko, S., Franceschi, F., Sternbach, H. & Nierhaus, K. H. Apparent association constants of tRNAs for the ribosomal A, P, and E sites. *J. Biol. Chem.* **267**, 4693–4702 (1992).



18. Sprinzl, M. & Richter, D. Free 3'-OH group of the terminal adenosine of the tRNA molecule is essential for the synthesis *in vitro* of guanosine tetraphosphate and pentaphosphate in a ribosomal system from *Escherichia coli*. *Eur. J. Biochem.* **71**, 171–176 (1976).
19. Friesen, J. D., Fiil, N. P., Parker, J. M. & Haseltine, W. A. A new relaxed mutant of *Escherichia coli* with an altered 50S ribosomal subunit. *Proc. Natl Acad. Sci. USA* **71**, 3465–3469 (1974).
20. Wendrich, T. M., Blaha, G., Wilson, D. N., Marahiel, M. A. & Nierhaus, K. H. Dissection of the mechanism for the stringent factor RelA. *Mol. Cell* **10**, 779–788 (2002).
21. Gropp, M., Strausz, Y., Gross, M. & Glaser, G. Regulation of *Escherichia coli* RelA requires oligomerization of the C-terminal domain. *J. Bacteriol.* **183**, 570–579 (2001).
22. Schreiber, G. *et al.* Overexpression of the relA gene in *Escherichia coli*. *J. Biol. Chem.* **266**, 3760–3767 (1991).
23. Valle, M. *et al.* Locking and unlocking of ribosomal motions. *Cell* **114**, 123–134 (2003).
24. Maris, C., Dominguez, C. & Allain, F. H. T. The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J.* **272**, 2118–2131 (2005).
25. Hogg, T., Mechold, U., Malke, H., Cashel, M. & Hilgenfeld, R. Conformational antagonism between opposing active sites in a bifunctional RelA/SpoT homolog modulates (p)ppGpp metabolism during the stringent response. *Cell* **117**, 57–68 (2004).
26. Yang, X. & Ishiguro, E. E. Dimerization of the RelA protein of *Escherichia coli*. *Biochem. Cell Biol.* **79**, 729–736 (2001).
27. English, B. P. *et al.* Single-molecule investigations of the stringent response machinery in living bacterial cells. *Proc. Natl Acad. Sci. USA* **108**, E365–E373 (2011).
28. Voorhees, R. M. & Ramakrishnan, V. Structural basis of the translational elongation cycle. *Annu. Rev. Biochem.* **82**, 203–236 (2013).
29. Li, W. *et al.* Effects of amino acid starvation on RelA diffusive behavior in live *Escherichia coli*. *Mol. Microbiol.* **99**, 571–585 (2016).
30. Nguyen, D. *et al.* Active starvation responses mediate antibiotic tolerance in biofilms and nutrient-limited bacteria. *Science* **334**, 982–986 (2011).

**Acknowledgements** We thank A. Kelley for providing tRNAs, A. Xu and J. Murray for their contributions to the early stages of this project, C. G. Savva for help with data collection, J. Grimmer and T. Darling for computing support, and X. Bai, G. Murshudov, S. H. W. Scheres, and S. Tan for discussions. This work was supported by grants to V.R. from the UK Medical Research Council (MC\_U105184332), the Wellcome Trust (WT096570), the Agouron Institute, and the Louis-Jeantet Foundation.

**Author Contributions** A.B., I.S.F., and V.R. designed the study. Y.G. purified ribosomes. I.S.F. prepared samples, optimized conditions and collected data. A.B. processed data and interpreted the cryo-electron microscopy reconstructions. A.B., I.S.F., and V.R. wrote the manuscript. All authors discussed and commented on the final manuscript.

**Author Information** The maps have been deposited with the EMDB under accession codes 8107–8115. Coordinates have been deposited with the Protein Data Bank under the accession code 5IQR. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to V.R. ([ramak@mrc-lmb.cam.ac.uk](mailto:ramak@mrc-lmb.cam.ac.uk)).

## METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

**RelA purification.** The full-length *E. coli* RelA protein containing an in-frame N-terminal octahistidine tag and the recognition signal for tobacco etch virus (TEV) protease was recombinantly expressed in *E. coli* BL21 cells, as previously described<sup>13</sup>. In a modification to the reported procedure, the protein was expressed for 3 h at 37°C after induction. Cells were collected by centrifugation at 4°C and lysed by sonication in ice-cold lysis buffer (50 mM Tris pH 7.5, 1 M KCl, 1 mM MgCl<sub>2</sub>, 15 mM imidazole, 2 mM DTT) containing 1× protease inhibitor cocktail (Roche). After centrifugation, the supernatant of the lysate was applied to a 5 ml HisTrap HP column (GE Healthcare) equilibrated in lysis buffer. After copious washing, a linear gradient of lysis buffer supplemented with 450 mM imidazole was used to elute the histidine-tagged RelA protein. RelA was dialysed against 2 l of buffer A (20 mM Tris pH 7.5, 0.6 M KCl, 1 mM MgCl<sub>2</sub> and 2 mM DTT) at 4°C for 2 h in the presence of TEV protease in a 1:200 mass ratio with RelA. A passage through a 5 ml HisTrap HP column equilibrated in buffer A was performed in a gradient of 0–450 mM imidazole to remove TEV protease and cleaved His tag. The early-eluting peak corresponding to digested RelA was used immediately to prepare complexes for freezing on cryo-EM grids to complete the procedure within a single day.

**Sample preparation for electron microscopy.** Ribosomes were purified from 50 g of *E. coli* MRE 600 cells grown to mid-log phase. Cells were disrupted by sonication in buffer R200 (20 mM Tris-HCl pH 7.5, 200 mM NH<sub>4</sub>Cl, 10 mM Mg(OAc)<sub>2</sub>, 6 mM β-mercaptoethanol, 0.1 mM benzimidazole, and 0.1 mM PMSF). Cell debris was removed by two rounds of centrifugation at 30,000g for 30 min. Ribosomes were pelleted through a 20 ml 1.1 M sucrose cushion using a 45 Ti rotor (Beckman Coulter) at 205,000g for 18 h. This process was repeated twice in buffer R200, and then once in buffer R500 (as R200, but with 500 mM NH<sub>4</sub>Cl). Between pelleting steps, pellets were dissolved in the appropriate buffer by gentle shaking for 2 h on ice. The final pellet was dissolved in buffer R60 (20 mM Tris pH 7.5, 60 mM NH<sub>4</sub>Cl, 10 mM Mg(OAc)<sub>2</sub>, 6 mM β-mercaptoethanol).

To remove non-ribosomal material and isolated subunits, ~208 OD<sub>260</sub> of sample were loaded on each of six 15–30% sucrose gradients in buffer R60 and centrifuged at 58,000g for 18 h with a SW28 rotor (Beckman Coulter). The peak fractions corresponding to 70S ribosomes were pooled and diluted to 180 ml with buffer G (5 mM HEPES pH 7.5, 50 mM KCl, 10 mM NH<sub>4</sub>Cl, 10 mM Mg(OAc)<sub>2</sub>, 6 mM β-mercaptoethanol), before a further pelleting step at 205,000g for 18 h. After washing with buffer G, the pellets were resuspended to a final concentration of ~6.8 μM in buffer G. Aliquots of 20 μl were flash-cooled in liquid nitrogen and stored at –80°C.

Aminoacylated fMet-tRNA<sup>fMet</sup> and uncharged tRNA<sup>Phe</sup> were produced as previously reported<sup>15</sup>. A modified version of Z4C mRNA<sup>15</sup> was chemically synthesized (GE Dharmacon) to include six codons for Phe (UUC) after the initiator fMet codon (AUG).

Complexes were formed by incubating 100 nM ribosome with the step-wise addition of 1 μM mRNA in buffer G100 (5 mM K-HEPES pH 7.0, 100 mM KCl, 10 mM NH<sub>4</sub>Cl, 10 mM Mg(OAc)<sub>2</sub>, 6 mM β-mercaptoethanol), 200 nM fMet-tRNA<sup>fMet</sup>, and 800 nM tRNA<sup>Phe</sup>. The complexes were incubated for 5 min at 37°C between each addition. Separately, 500 nM RelA was pre-incubated with 1 μM AMPCPP and 1 μM GDP for 10 min at 37°C, and then added to the ribosome complex to a final volume of 100 μl. To stabilize the codon–anticodon interaction of the A-site tRNA, 1 μM paromomycin was added. After a further 5 min incubation, the sample was cooled to 4°C and used immediately to prepare grids for electron microscopy. **Electron microscopy.** Aliquots of 3 μl of the RelA complex were incubated for 20 s on glow-discharged holey carbon grids (Quantifoil R2/2), on which a ~30 Å-thick custom-made amorphous carbon film had previously been deposited. The grids were blotted for 5 s in 100% humidity at 4°C before being flash-cooled in liquid ethane using a Vitrobot MKII (FEI).

Grids were transferred to a Polara G2 microscope (FEI) operated at 300 kV. Images were recorded with the EPU automated data acquisition software on a Falcon III direct electron detector (FEI) at a calibrated magnification of 104,478 (yielding a pixel size of 1.34 Å). Images were collected with a total dose of 35 e<sup>−</sup> per Å<sup>2</sup> and a defocus range of –1.8 to –3.0 μm. A bespoke system was used to intercept the videos from the detector at a speed of 30 frames for the 1.1-s exposures.

**Image processing.** All micrographs that showed signs of astigmatism, substantial drift, or poor ice were discarded. The frames of the remaining micrographs were aligned using whole-image motion correction<sup>31</sup> to reduce beam-induced blurring. Parameters of the contrast transfer function for each motion-corrected micrograph were obtained using Gctf<sup>32</sup>. The interactive semi-automatic swarm tool in the e2boxer.py program of EMAN2<sup>33</sup> was used to select particles from the images.

We used reference-free two-dimensional class averaging in RELION<sup>34</sup> to discard non-ribosomal particles and particles of isolated subunits. After two-dimensional classification, Euler angles for each particle were assigned using three-dimensional refinement in RELION<sup>34</sup> using a 60 Å low-pass-filtered reconstruction of EMDB-2373<sup>13</sup> as an initial reference.

This resulted in a consensus reconstruction in which density for tRNAs and RelA were apparent, but at lower occupancy than the ribosome. To enrich for particles containing RelA, we employed focused classification with signal subtraction (FCwSS)<sup>12</sup> using a mask over the A-site tRNA and RelA. This yielded a class in which the A site was fully occupied. The movements of particles within this class were corrected in RELION and the contributions of each frame weighted using a resolution-dependent radiation damage model<sup>35</sup> to generate ‘polished particles’. Polished particles from different data sets were commingled and further classified and refined to yield a final data set of 164,353 particles with a nominal resolution of 3.0 Å (Extended Data Table 1). Owing to on-ribosome conformational heterogeneity, subsequent rounds of FCwSS were used to isolate different states of RelA. Prior to visualization, all density maps were corrected for the modulation transfer function (MTF), and then sharpened by applying a B-factor that was estimated using automated procedures<sup>36</sup>. Local resolution was quantified with ResMap<sup>37</sup> (Extended Data Fig. 2).

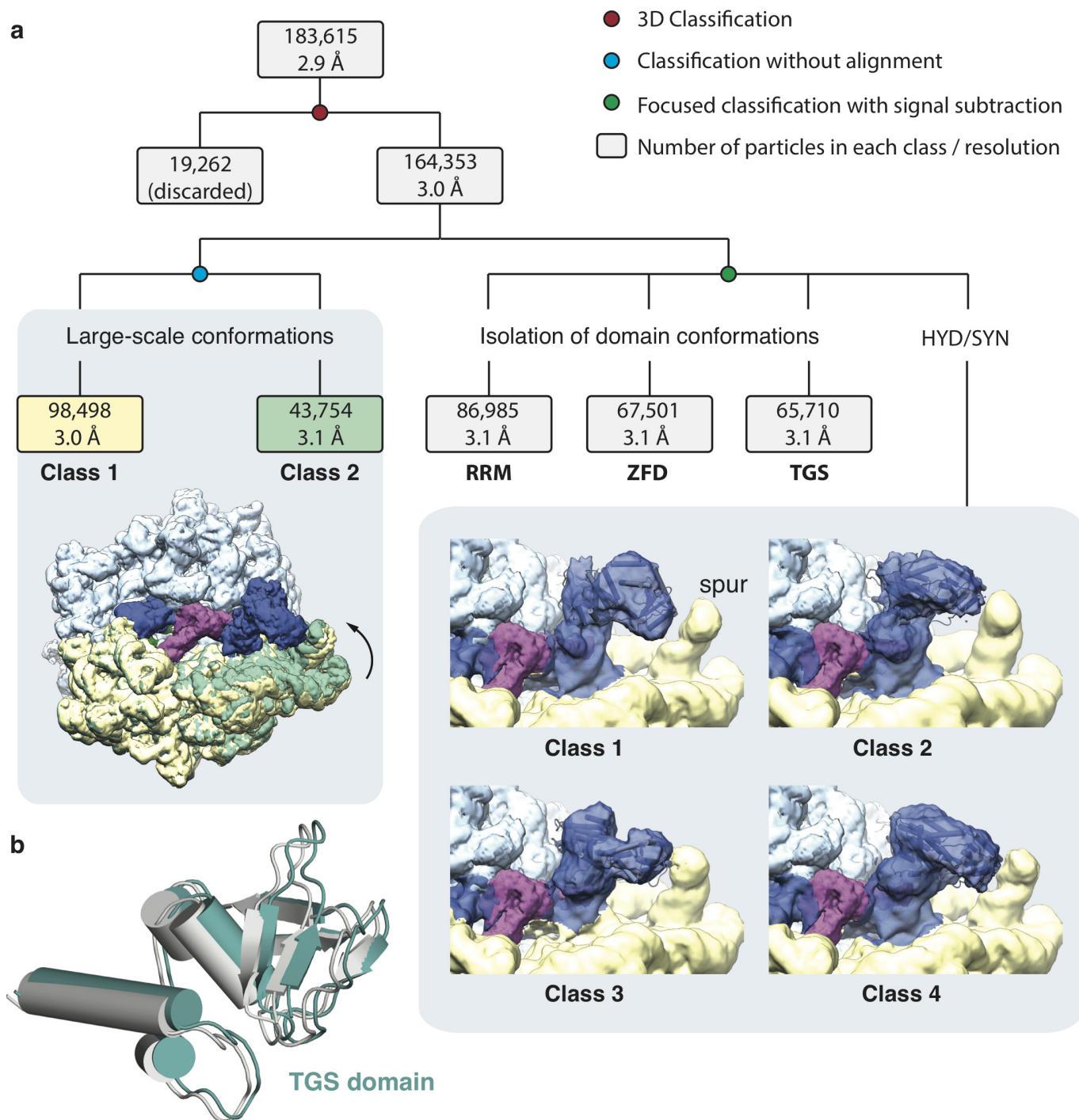
**Model building.** The reconstruction was initially interpreted by docking the high-resolution crystal structure of the *Escherichia coli* 70S ribosome (PDB accession code 4YBB<sup>38</sup>) into the map using Chimera<sup>39</sup>. Models for uL10, bL9, and bL31 were taken from the cryo-EM reconstruction of *E. coli* 70S in complex with EF-Tu (PDB accession code 5AFI<sup>40</sup>). Initial models for the tRNAs and mRNA were taken from the *Thermus thermophilus* 70S in complex with EF-Tu (PDB accession code 4V5L<sup>41</sup>). Density for paromomycin was clearly discernible within rRNA helix 44 of the small subunit, and fitted with a model. Unlike previous crystal structures<sup>42</sup>, we do not observe a second paromomycin binding site within the large subunit.

Homology models generated using I-TASSER<sup>43</sup> were used to guide model building for *E. coli* RelA. A model for the hydrolase and synthetase domains was obtained from the crystal structure of the bifunctional catalytic domain from a *Streptococcus equisimilis* RSH<sup>25</sup> (PDB accession code 1VJ7). A model for the RRM was generated using the structure of the RRM domain from a *Chlorobium tepidum* RSH (PDB accession code 3IBW). A model for the TGS domain was derived from the crystal structure of the TGS domain from a *Clostridium leptum* RSH (PDB accession code 3HVZ). The ZFD and connecting elements were built *de novo* (Extended Data Table 2). The fit of all models to the map was optimized using real space refinement in Coot<sup>44</sup>.

**Model refinement and validation.** Reciprocal space refinement was carried out in REFMAC v5.8 optimized for EM maps using external restraints generated by ProSMART and LIBG<sup>44</sup>. Fourier shell correlation (FSC)<sub>average</sub> was monitored during refinement and the final model was validated using MolProbity<sup>45</sup> (Extended Data Table 1). Cross-validation against over-fitting was performed as previously described<sup>44,46</sup>. Figures were generated using PyMOL<sup>47</sup>, Chimera<sup>39</sup>, or Coot/Raster3D<sup>48,49</sup>.

- Li, X. *et al.* Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* **10**, 584–590 (2013).
- Zhang, K. Gctf: real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2015).
- Tang, G. *et al.* EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
- Scheres, S. H. W. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
- Scheres, S. H. Beam-induced motion correction for sub-megadalton cryo-EM particles. *elife* **3**, e03665 (2014).
- Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
- Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–65 (2013).
- Noeske, J. *et al.* High-resolution structure of the *Escherichia coli* ribosome. *Nat. Struct. Mol. Biol.* **22**, 336–341 (2015).
- Pettersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
- Fischer, N. *et al.* Structure of the *E. coli* ribosome-EF-Tu complex at <3 Å resolution by Cs-corrected cryo-EM. *Nature* **520**, 567–570 (2015).
- Voorhees, R. M., Schmeing, T. M., Kelley, A. C. & Ramakrishnan, V. The mechanism for activation of GTP hydrolysis on the ribosome. *Science* **330**, 835–838 (2010).
- Borovinskaya, M. A. *et al.* Structural basis for aminoglycoside inhibition of bacterial ribosome recycling. *Nat. Struct. Mol. Biol.* **14**, 727–732 (2007).

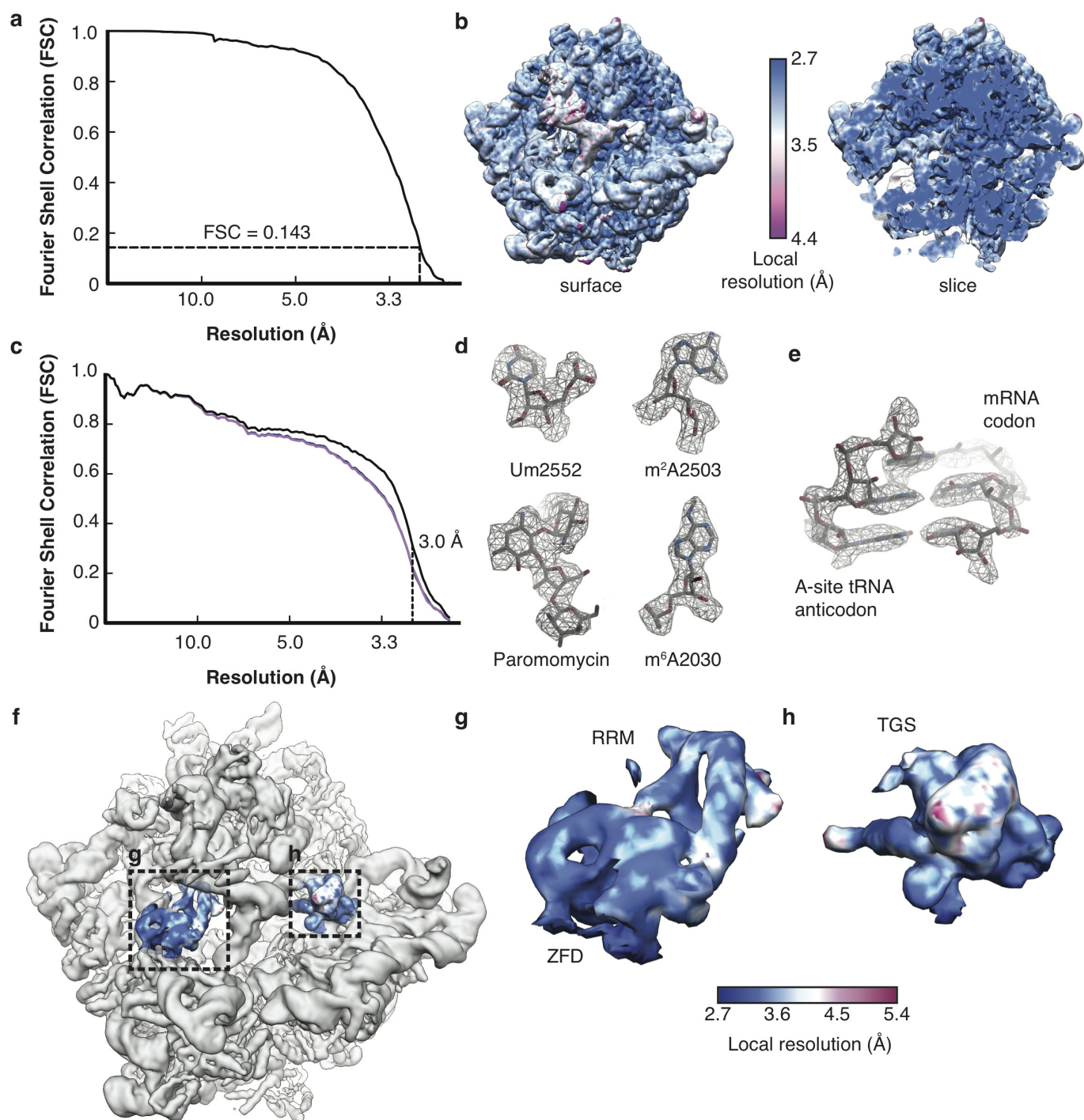
43. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, 40 (2008).
44. Brown, A. *et al.* Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta Crystallogr. D* **71**, 136–153 (2015).
45. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2009).
46. Amunts, A. *et al.* Structure of the yeast mitochondrial large ribosomal subunit. *Science* **343**, 1485–1489 (2014).
47. DeLano, W. L. The PyMOL molecular graphics system. (2002).
48. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
49. Merritt, E. A. & Bacon, D. J. Raster3D: photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524 (1997).
50. Stivala, A., Wybrow, M., Wirth, A., Whisstock, J. C. & Stuckey, P. J. Automatic generation of protein structure cartoons with Pro-origami. *Bioinformatics* **27**, 3315–3316 (2011).
51. Baker, N. A., Sept, D., Joseph, S. & Holst, M. J. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl Acad. Sci. USA* **98**, 10037–10041 (2001).



**Extended Data Figure 1 | *In silico* 3D classification scheme.** **a**, All particles were subjected to 2D classification, from which non-ribosomal particles were discarded, before 3D refinement. To isolate particles containing A-site tRNA and RelA, 3D classification focused on occupancy of the ribosomal A site was performed. Refinement of these 183,615 particles resulted in a reconstruction with a nominal resolution of 2.9 Å. A second round of 3D classification isolated 164,353 well-aligned particles. Conformational heterogeneity of the ribosome was resolved by 3D classification without alignment, which identified two dominant classes in which the body of the small subunit occupies different positions (indicated with an arrow). Class 1 was used as the reference for model building, refinement, and interpretation. To resolve additional conformational

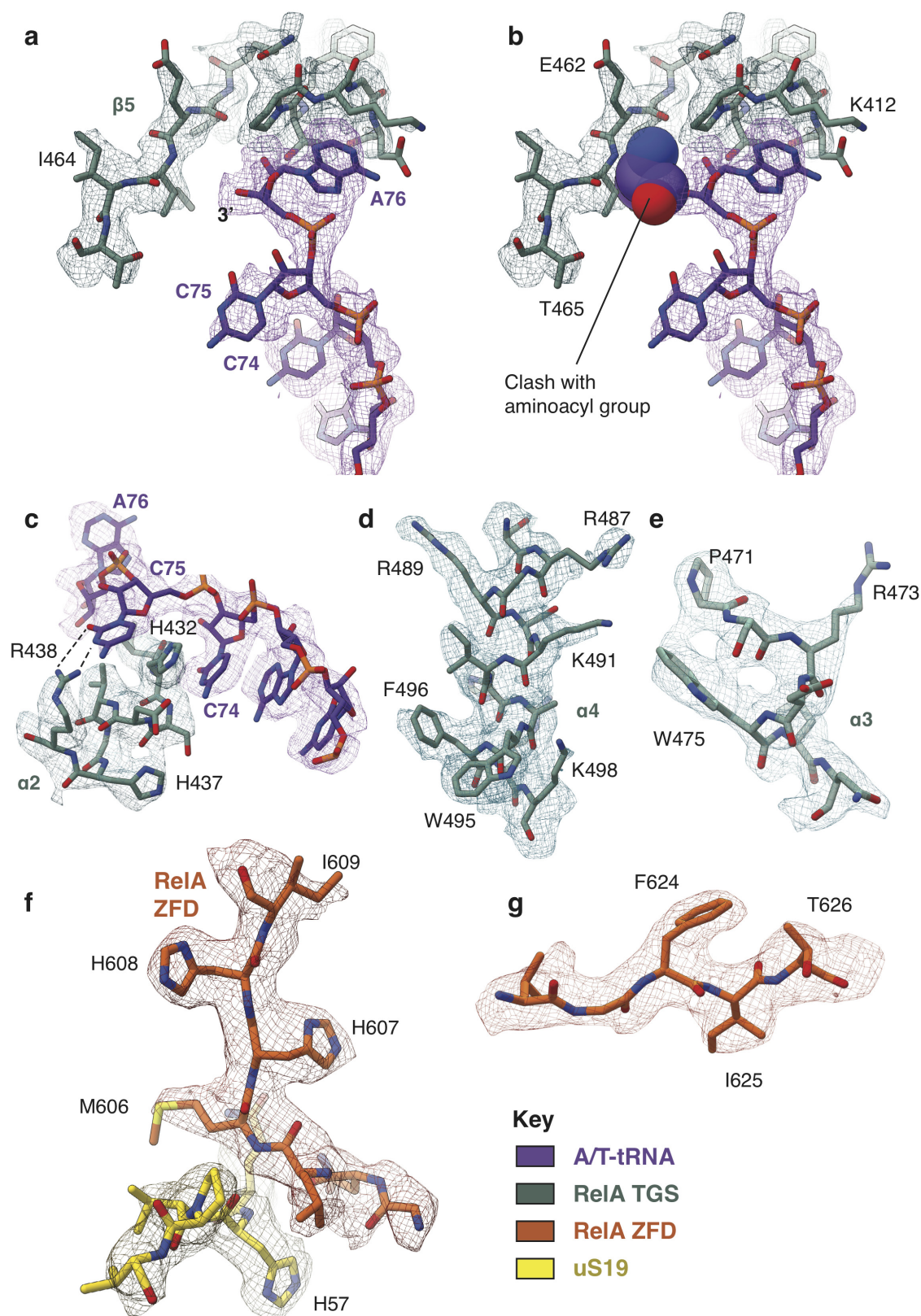
heterogeneity of RelA, focused classification with signal subtraction (FCwSS) was performed on each domain, with the hydrolase (HYD) and synthetase (SYN) domains treated as a single unit. For the RRM, zinc-finger and TGS domains a single class was isolated in which the density was better resolved than in the reference class. The overall resolution of the reconstructions are reported according to the Fourier shell correlation (FSC) = 0.143 criterion. Multiple conformations of the HYD and SYN domains were identified, with the four best-resolved classes shown. Together these account for 42% of the particles. **b**, The two main classes for the TGS domain provide an example of the small conformational differences that were isolated using FCwSS.





**Extended Data Figure 2 | Quality of maps and models.** **a**, FSC curve for the EM map. **b**, The unfiltered and unsharpened density map, in both surface and slice view, coloured by local resolution. **c**, Fit of models to maps. FSC curves calculated between the refined model and the final map (black), with the self- and cross-validated correlations in blue and magenta, respectively. Information beyond 3.0 Å was not used during refinement and preserved for validation. **d**, **e**, Examples of high-resolution features of the map. **d**, Density for selected rRNA modifications and

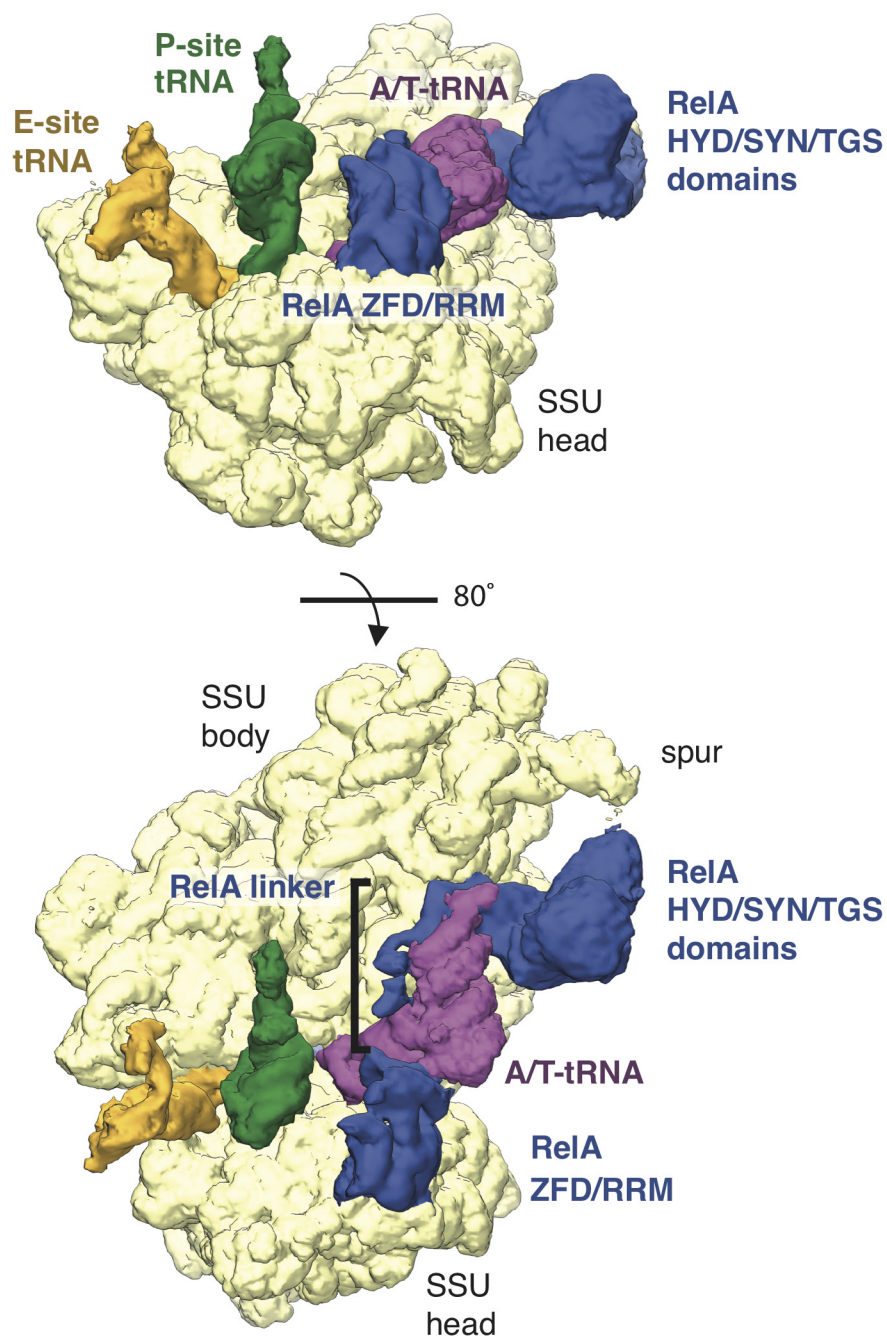
paromomycin. **e**, Density for the codon–anticodon interaction in the A site. **e**, Unfiltered and unsharpened map of RelA bound to the bacterial ribosome, showing the ribosome-binding RelA domains coloured by local resolution according to the FCwSS maps (see Extended Data Fig. 1). The regions amplified in panels **f** and **g** are highlighted. **f**, The RelA ZFD and RRM coloured by local resolution. **g**, The TGS domain coloured by local resolution.



**Extended Data Figure 3 | Examples of RelA density.** **a**, Density for the interaction between the 3' CCA of the A/T-tRNA and the TGS domain. **b**, A modelled tRNA<sup>Ala</sup> demonstrates that even the smallest aminoacyl groups would clash with RelA. The sphere size of the atoms of the aminoacyl group corresponds to their van der Waals radii. **c**, C74 stacks

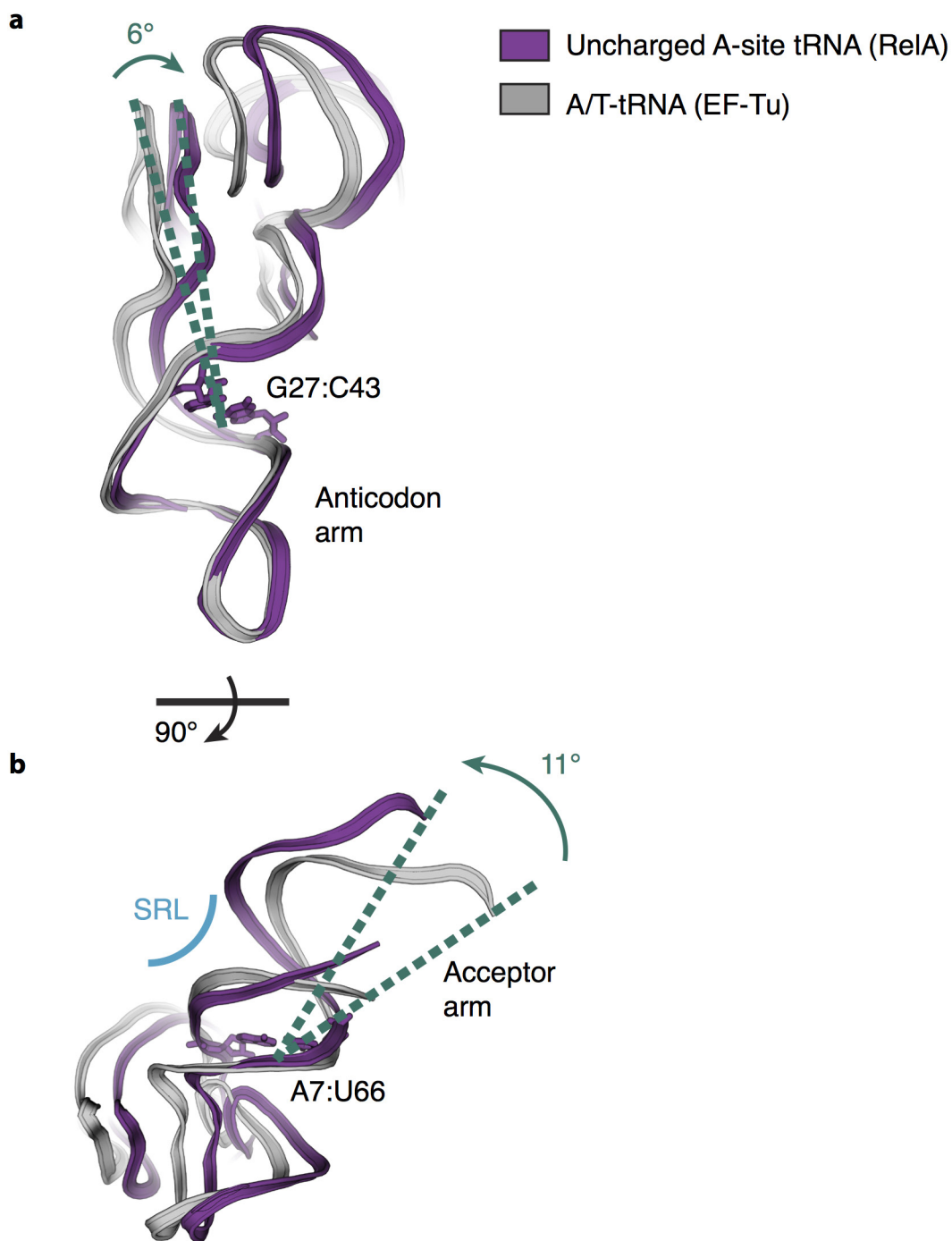
with His432 and C75 can potentially interact with Arg438 of the TGS domain. **d**, Density for helix  $\alpha 4$  of the TGS domain. **e**, Density for TGS  $\alpha 3$ . **f**, Density for the interaction between the ZFD and uS19, showing distinctive density for two consecutive histidine residues. **g**, Example of side chain density used for the *de novo* building of the ZFD.





**Extended Data Figure 4 | RelA domains are connected by flexible linkers.** Two related views showing the density that connects the RelA HYD, SYN and TGS domains with the ZFD/RRM. The linker runs between the A/T-tRNA and the ribosome, but remains flexible as

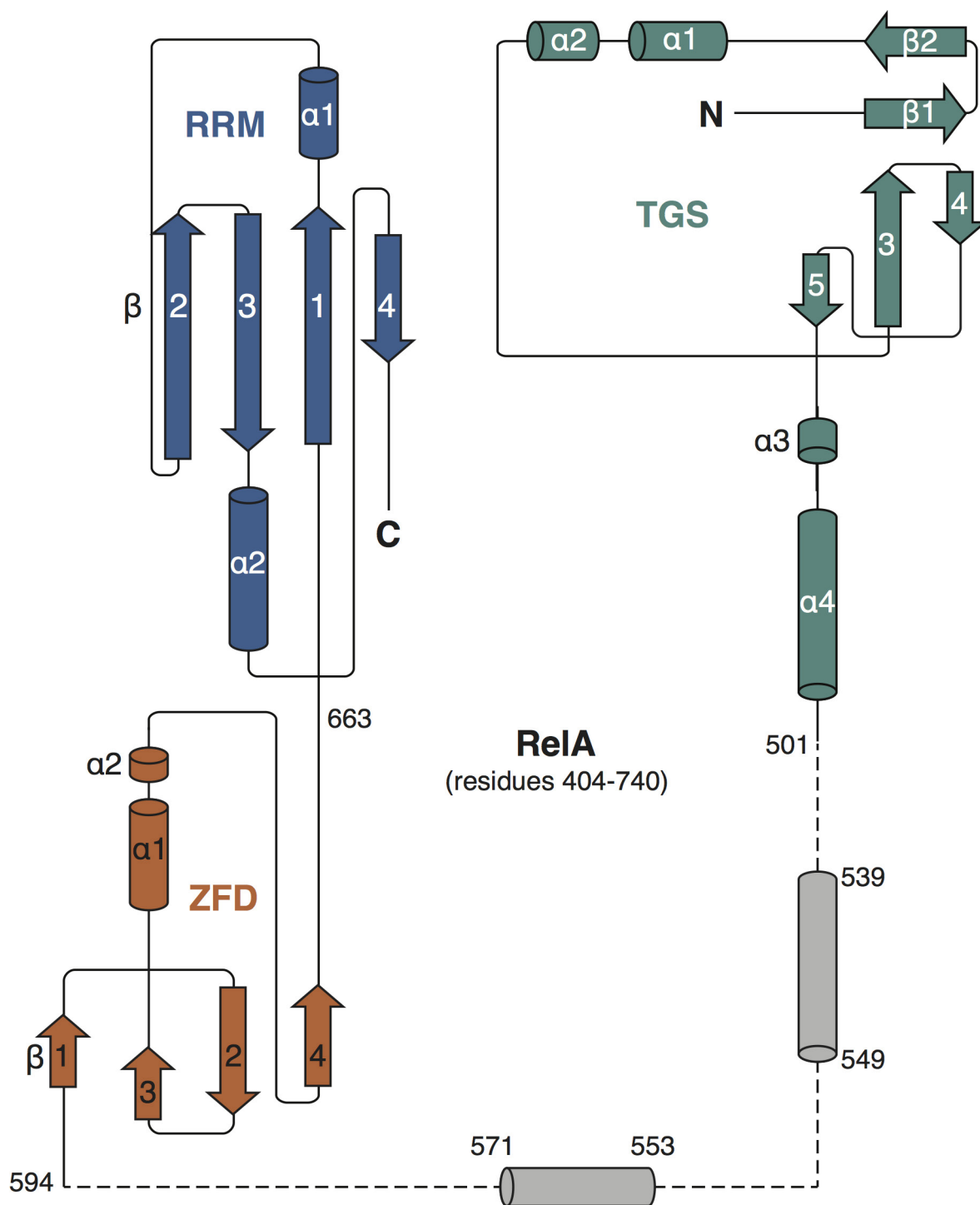
suggested by the weak and broken density. The figure shows the unfiltered, unsharpened density map for the ribosomal small subunit (SSU) of class 4 (Extended Data Fig. 1) with the large subunit removed for clarity.



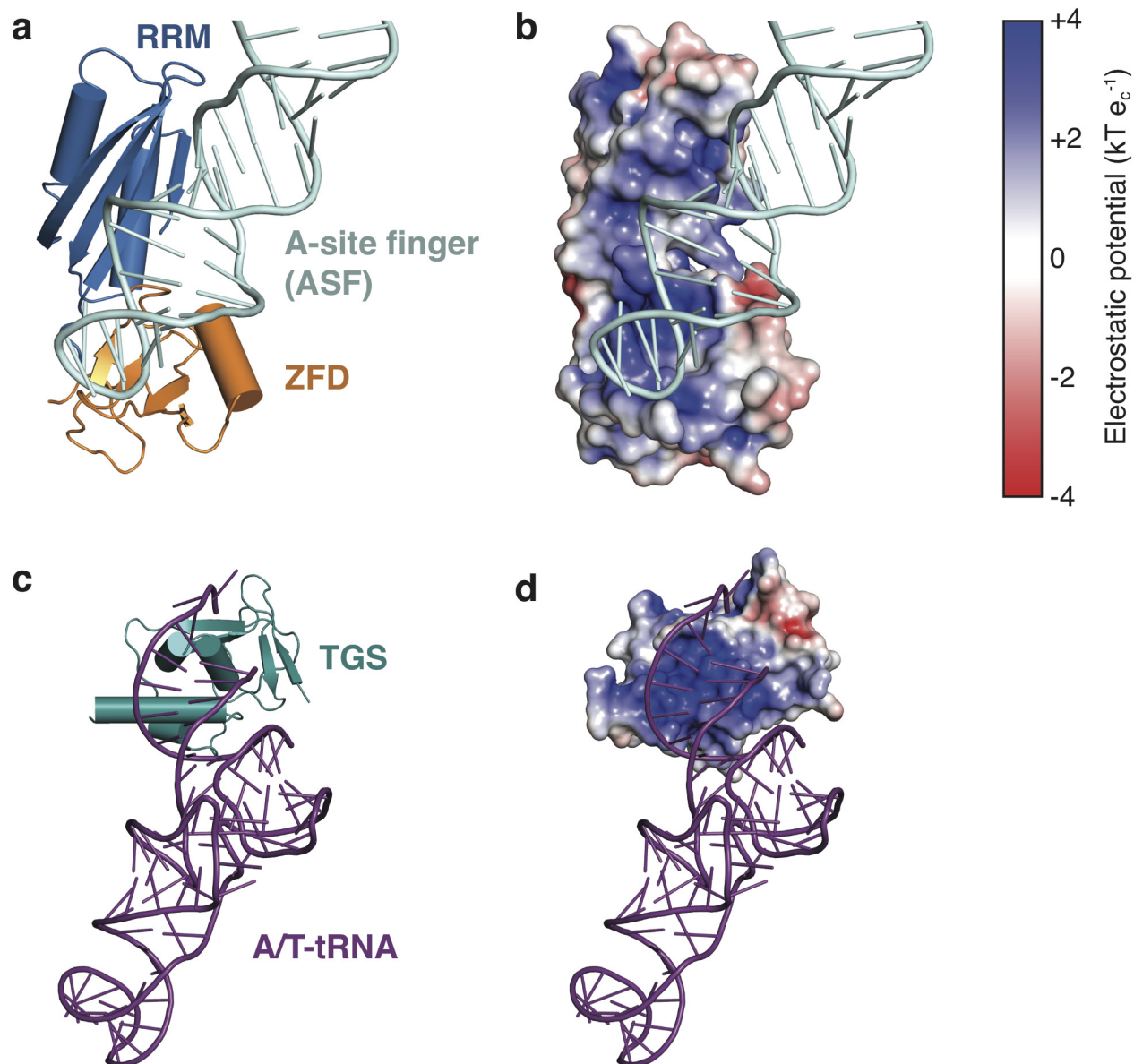
**Extended Data Figure 5 | The conformation of uncharged A-site tRNA in the presence of RelA is distinct from aminoacylated A/T-tRNA in the presence of EF-Tu. a,** The ASLs of A-site tRNA (purple) and A/T-tRNA (grey) superpose until base-pair 27:43. At this point, the A-site tRNA is

distorted so that the tRNA elbow regions are separated by a 6° rotation. **b,** A second 11° rotation occurs at base-pair 7:66 of the acceptor stem so that the A-site tRNA in the presence of RelA is closer to the ribosomal SRL.



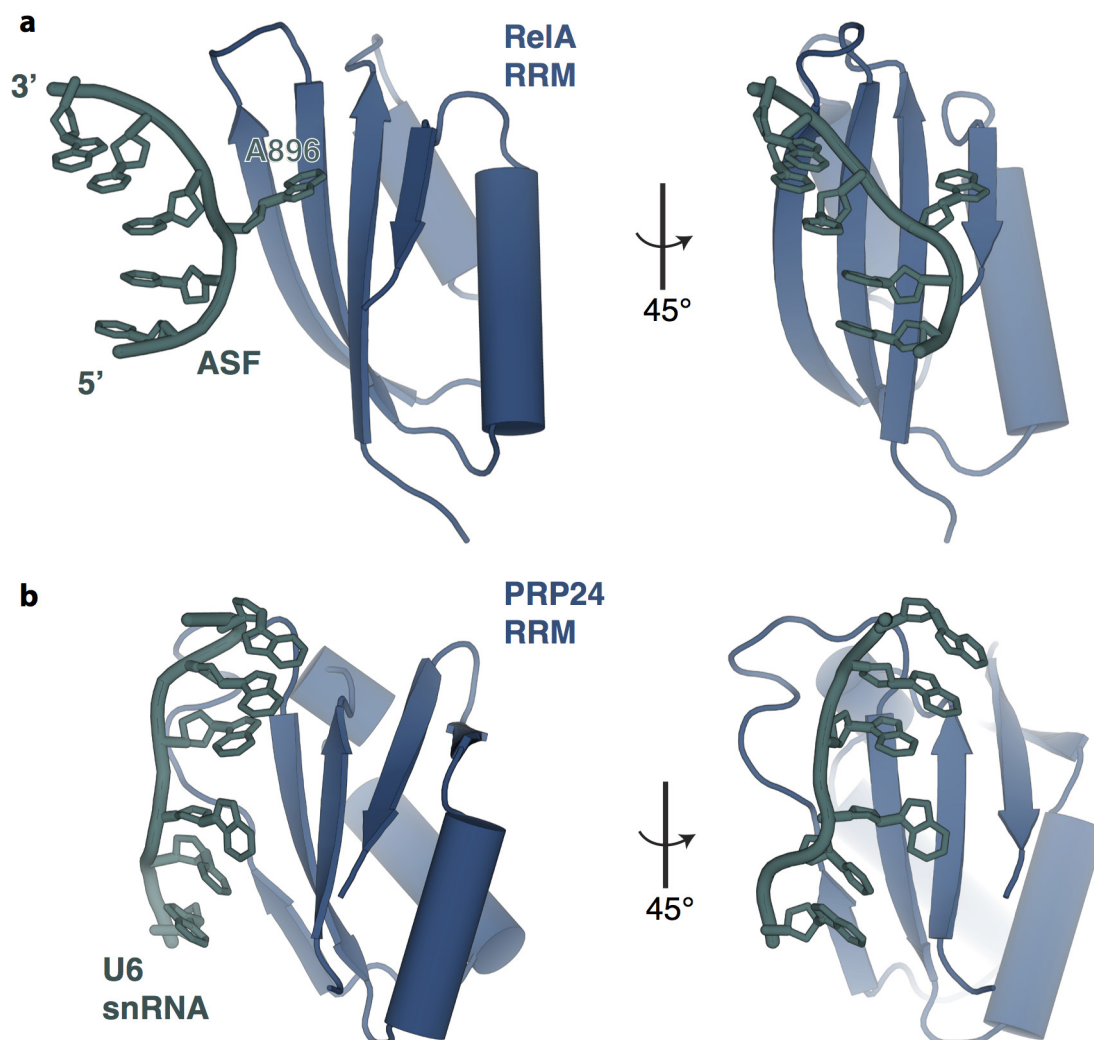


**Extended Data Figure 6 | RelA topology diagram.** Secondary structure elements for RelA residues 404–740 are numbered separately for each domain. Unbuilt sections are shown as dashed lines. Topologies were extracted using Pro-origami<sup>50</sup>.



**Extended Data Figure 7 | RelA binds RNA through electropositive surfaces.** **a**, The ZFD and RRM of RelA act together to recognize the ASF of the LSU rRNA. **b**, As in **a**, but with the ZFD and RRM shown in surface representation coloured by electrostatic potential. **c**, The RelA TGS domain binds the acceptor arm of the A/T-tRNA. **d**, As in **c**, but

with the TGS domain in surface representation coloured by electrostatic potential. Electrostatic potentials were calculated using APBS<sup>51</sup>, where  $k$  is Boltzmann's constant,  $T$  is the temperature of the calculation (310 K) and  $e_c$  is the charge of an electron.



**Extended Data Figure 8 | RelA contains an RNA recognition motif (RRM).** **a**, The RRM from RelA binds the ASF (nucleotides 894–899 shown) through the face of the  $\beta$ -sheet. **b**, RRM s recognize a wide variety

of RNA molecules, but share a common fold and a similar protein–RNA interface, for example in the interaction between PRP24 and U6 small nuclear RNA (PDB accession code 4N0T).

Extended Data Table 1 | Refinement and model statistics

**Data Collection**

Particles	98,498
Pixel size (Å)	1.34
Defocus mean (μm)	2.0
Defocus range (μm)	0.5 - 3.9
Voltage (kV)	300
Electron dose (e <sup>-</sup> Å <sup>-2</sup> )	35

**Model composition**

Non-hydrogen atoms	151,800
Protein residues	6,229
RNA bases	4,782
Ligands (Zn <sup>2+</sup> /Mg <sup>2+</sup> /H <sub>2</sub> O/PAR)	3/312/2/1

**Refinement**

Resolution (Å)	3.0
Map sharpening B-factor (Å <sup>2</sup> )	-40
Average B factor (Å <sup>2</sup> )	106.0
FSC <sub>average</sub>	0.82

**Rms deviations**

Bond lengths (Å)	0.005
Bond angles (°)	1.1

**Validation (proteins)**

Molprobity score	2.5 (95 <sup>th</sup> )
Clashscore, all atoms	6.5 (100 <sup>th</sup> )
Good rotamers (%)	94.2

**Ramachandran plot**

Favored (%)	91.3
Outliers (%)	1.3

**Validation (RNA)**

Correct sugar puckers (%)	99.0
Good backbone conformations (%)	78.4

The RelA HYD/SYN domain (residues 15–353) was modelled as a rigid-body fitted homology model and excluded from refinement.



Extended Data Table 2 | Modelled residues

Residues	Domain	Model details
15-353	HYD/SYN	Rigid body fitted homology model based on PDB accession code 1VJ7.
404-501	TGS	Model based on PDB accession code 3HVZ adjusted to fit the density.
539-549	Linker	Modeled as a poly(alanine) helix.
553-571	Linker	Modeled as a poly(alanine) helix.
594-663	ZFD	Built <i>de novo</i> .
664-740	RRM	Model based on PDB accession code 3IBW adjusted to fit the density.

# Charge–density analysis of an iron–sulfur protein at an ultra–high resolution of 0.48 Å

Yu Hirano<sup>1\*†</sup>, Kazuki Takeda<sup>1\*</sup> & Kunio Miki<sup>1</sup>

The fine structures of proteins, such as the positions of hydrogen atoms, distributions of valence electrons and orientations of bound waters, are critical factors for determining the dynamic and chemical properties of proteins. Such information cannot be obtained by conventional protein X-ray analyses at 3.0–1.5 Å resolution, in which amino acids are fitted into atomically unresolved electron-density maps and refinement calculations are performed under strong restraints<sup>1,2</sup>. Therefore, we usually supplement the information on hydrogen atoms and valence electrons in proteins with pre-existing common knowledge obtained by chemistry in small molecules. However, even now, computational calculation of such information with quantum chemistry also tends to be difficult, especially for polynuclear metalloproteins<sup>3</sup>. Here we report a charge-density analysis of the high-potential iron–sulfur protein from the thermophilic purple bacterium *Thermochromatium tepidum* using X-ray data at an ultra-high resolution of 0.48 Å. Residual electron densities in the conventional refinement are assigned as valence electrons in the multipolar refinement. Iron 3d and sulfur 3p electron densities of the Fe<sub>4</sub>S<sub>4</sub> cluster are visualized around the atoms. Such information provides the most detailed view of the valence electrons of the metal complex in the protein. The asymmetry of the iron–sulfur cluster and the protein environment suggests the structural basis of charge storing on electron transfer. Our charge-density analysis reveals many fine features around the metal complex for the first time, and will enable further theoretical and experimental studies of metalloproteins.

Fine structural information, including the positions of hydrogen atoms and distributions of valence electrons, is essential for understanding the full properties of protein molecules; in particular, of electron transfer metalloproteins in photosynthesis and respiration. However, it is difficult to obtain such information with theoretical and experimental analyses. Therefore, experimental visualization of valence electrons and hydrogen atoms is of great value in elucidating the molecular mechanisms of protein functions. We have previously performed crystallographic investigations of photosynthetic proteins from *T. tepidum*<sup>4–6</sup>. The high-potential iron–sulfur protein (HiPIP) is critical for electron transfer in bacterial photosynthesis. HiPIP from *T. tepidum* consists of 83 residues and possesses one Fe<sub>4</sub>S<sub>4</sub> cluster at the centre of the protein. The protein gives high-quality crystals suitable for X-ray crystallographic analysis at ultra-high resolution<sup>7</sup>. It has been reported that the redox properties of the Fe<sub>4</sub>S<sub>4</sub> cluster are attributable to interactions between the cluster and surrounding ligands, both in Fe<sub>4</sub>S<sub>4</sub> and iron–sulfur proteins<sup>8,9</sup> as well as in analogue compounds<sup>10</sup>. Therefore, we performed an experimental charge-density analysis of HiPIP in its reduced state. This is the first case, to our knowledge, of charge-density analysis being applied to metalloproteins, although it has been used previously for several proteins<sup>11–14</sup>.

A diffraction data set at 0.48 Å resolution was collected using high-energy synchrotron X-rays ( $\lambda = 0.45$  Å) (Extended Data Table 1

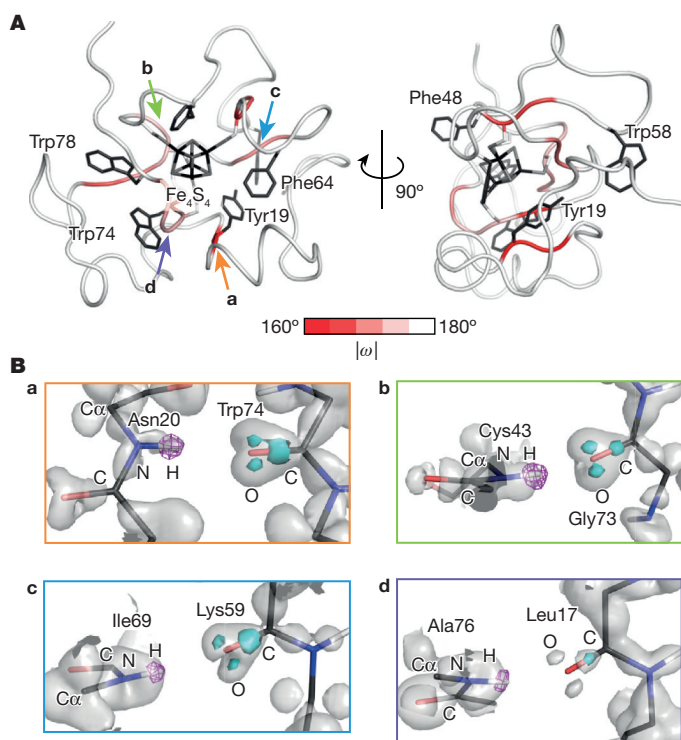
and Extended Data Fig. 1). This is equivalent to one of the highest-resolution data sets in the Protein Data Bank (<http://www.rcsb.org/pdb/home/home.do>), and it allowed us to perform structure refinement using the conventional independent spherical atom model (ISAM) without geometric restraints. Even after ISAM refinement, however, many residual electron densities were observed around each atom (Extended Data Fig. 2). Peaks of residual electron density were on the covalent bonds and around the carbonyl oxygen atoms of peptide bonds (Extended Data Fig. 2a), as well as on the covalent bonds of aromatic rings (Extended Data Fig. 2b). In the Fe<sub>4</sub>S<sub>4</sub> cluster, peaks were symmetrically distributed around the Fe atoms (Extended Data Fig. 2c). The charge-density information of the electron densities was analysed with a multipolar atom model (MAM)<sup>15</sup>. The *R* factor was reduced from 8.24% (*R*<sub>free</sub> = 8.63%) to 7.16% (*R*<sub>free</sub> = 7.80%) by applying the charge-density analysis. The final structure contained hydrogen atoms of all 83 residues. In addition, 42 hydrogen atoms of water molecules were also included (Extended Data Table 1). The deformation map<sup>15</sup> reveals the distribution of valence electrons in the protein. The fine structural information both of hydrogen atoms and of valence electrons highlights detailed views of intra-molecular interactions in the protein.

The dihedral angle  $\omega$  of C <sub>$\alpha$</sub> –C–N'–C <sub>$\alpha$</sub> ' in a peptide bond, where prime symbols represent the atoms in the next residue, defines the planarity of the peptide bond, and the planar *trans* peptide shows an angle  $\omega$  of 180°. Although some non-planar peptide bonds have been observed in protein crystal structures determined at ultra-high resolutions<sup>16</sup>, a planar geometry is adopted in the protein structure at ordinary resolutions<sup>17</sup>. In HiPIP, the peptides deviating approximately 10° or more from the planar  $\omega$  angle are mainly located around the cysteine residues that are covalently bound to the Fe atoms of the Fe<sub>4</sub>S<sub>4</sub> cluster (Fig. 1A). The distortion of peptide bonds is also observed in the positions of amide hydrogen atoms (Fig. 1B). Almost all hydrogen atoms deviating from the C–N'–C <sub>$\alpha$</sub> ' plane are in the non-planar peptide bonds (Extended Data Table 2). These hydrogen atoms are concentrated in the proximal region of the Fe<sub>4</sub>S<sub>4</sub> cluster. The deviation of hydrogen atoms is plausibly influenced by the hydrogen bonds they form, in which the donor hydrogen atoms are pointed towards the lone pair electron density of the acceptor atoms (Fig. 1B). The same features have also been observed for the donor hydrogen atoms in the crystal structures of small molecules<sup>18</sup>. Three (Cys43, Cys61 and Cys75) out of four cysteine residues have distorted peptide bonds, while Cys46 does not (Extended Data Table 2). NMR data suggest that  $\sigma$ -type delocalization of Fe orbitals occurs through Fe–(Cys–S <sub>$\gamma$</sub> ) bonds<sup>19</sup>. This may further affect distortion of the peptide bonds. Indeed, Cys46 has the longest Fe–(Cys–S <sub>$\gamma$</sub> ) bond (Extended Data Table 3). The non-planarity causes partial breakdown of the resonance of the peptide bonds<sup>20,21</sup>.

Non-spherical distribution of electron density is clearly observed around the atoms of Fe<sub>4</sub>S<sub>4</sub>(Cys–S <sub>$\gamma$</sub> )<sub>4</sub> (Fig. 2a). Lobes of electron density

<sup>1</sup>Department of Chemistry, Graduate School of Science, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan. <sup>†</sup>Present address: Quantum Beam Science Research Directorate, National Institutes for Quantum and Radiological Science and Technology, Tokai-mura, Ibaraki 319-1195, Japan.

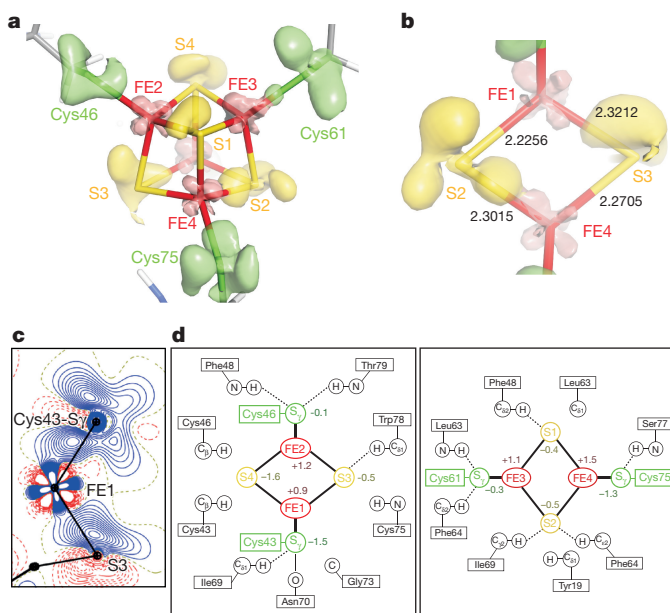
\*These authors contributed equally to this work.



**Figure 1 | Crystal structure of HiPIP at 0.48 Å resolution.** **A**, Overall structure of HiPIP. Non-hydrogen atoms of aromatic side-chains, non-hydrogen atoms of cysteine residues and atoms of the  $\text{Fe}_4\text{S}_4$  cluster are depicted as stick models. The main-chain atoms are shown as a tube model and single conformational residues are coloured according to the omega angles ( $|\omega|$ ) from 160° (red) to 180° (white). The main-chain atoms of multi-conformational residues are coloured white. The aromatic side-chains are coloured dark grey. The Fe and S atoms of  $\text{Fe}_4\text{S}_4(\text{Cys-S}_\gamma)_4$  are coloured black and light grey, respectively. The orange, green, light blue and violet arrows represent the peptide bonds indicated in **B**. **B**, Hydrogen atoms in the peptide bonds of (a) Asn20, (b) Cys43, (c) Ile69 and (d) Ala76. The static deformation maps are shown as grey and cyan surfaces contoured at the levels of +0.1 and +0.3 electrons per cubic Å, respectively. The omit maps of amide hydrogens are shown as a pink mesh contoured at the  $3.0\sigma$  level.

are symmetrically distributed around the Fe atoms, and correspond to the distribution of the Fe 3*d*-orbital electrons. The shapes of these electron-density lobes are quite similar among the four Fe atoms. In contrast to the Fe atoms, the bridging S and Cys-S $\gamma$  atoms are surrounded by more diffuse electron density that corresponds to the distribution of the S 3*p*-orbital electrons. The deformation electron densities of the individual S atoms do not exhibit any apparent similarity, unlike in the case of the Fe atoms. In some Fe–S bonds, the Fe 3*d* electron density points towards the S 3*p* electron density (Fig. 2b, c). Similar interactions in the valence density have been reported for the Ni–N bond in a transition metal complex<sup>22</sup>. However, the 3*d*–3*p* overlaps are smaller for bonds with shorter Fe–S bonds such as FE1–S2, while they are larger for longer bonds such as FE1–S3, in general.

The charge-density analysis presents the atomic charges of  $\text{Fe}_4\text{S}_4(\text{Cys-S}_\gamma)_4$  in HiPIP (Fig. 2d and Extended Data Table 4). The total atomic charge of  $\text{Fe}_4\text{S}_4(\text{Cys-S}_\gamma)_4$  is –1.5. This value is close to the total of the formal charges of  $\text{Fe}_4\text{S}_4(\text{Cys-S}_\gamma)_4$  in the reduced state. The atomic charge of FE1 (+0.9) is lower than those of FE2 (+1.2), FE3 (+1.1) and FE4 (+1.5). The atomic charge of S4 (–1.6) is significantly lower than those of other bridging S atoms (from –0.5 to –0.4). The atomic charges of Cys43-S $\gamma$  (–1.5) and Cys75-S $\gamma$  (–1.3) are lower than those of Cys46-S $\gamma$  (–0.1) and Cys61-S $\gamma$  (–0.3). The FE1 atom is coordinated by the ligand sulfur atoms, with the sum of the charges (–4.1) being more negative than for the other three Fe atoms (–2.7 to –2.6).

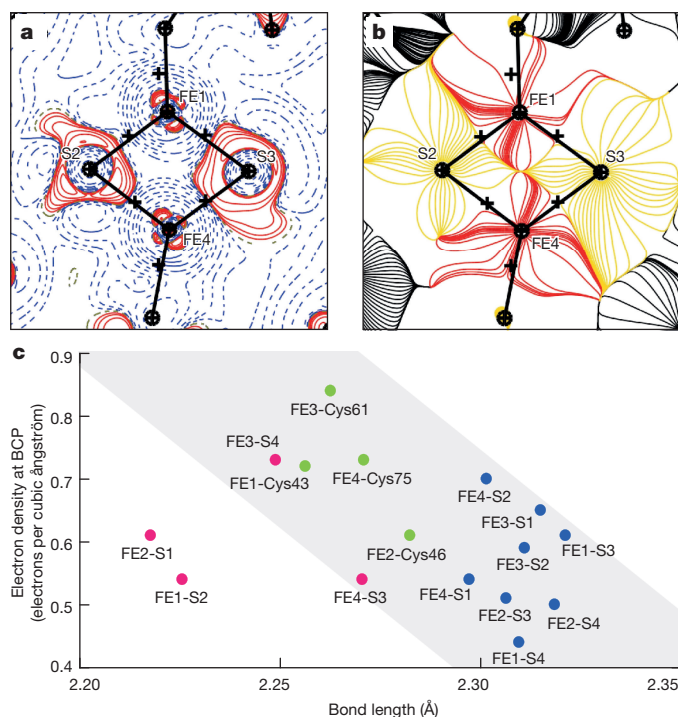


**Figure 2 | The  $\text{Fe}_4\text{S}_4$  cluster.** **a**, The three-dimensional representation of the static deformation map of  $\text{Fe}_4\text{S}_4(\text{Cys-S}_\gamma)_4$ . The orientation of the cluster is the same as the left panel of Fig. 1A. The isosurface represents the electron density contoured at the level of +0.2 electrons per cubic Å. **b**, A close-up view for the plane of FE1–S3–FE4–S2. Bond lengths are indicated in the figure. **c**, The static deformation maps of the plane consisting of FE1, S3 and Cys43-S $\gamma$  atoms. The contour interval is 0.05 electrons per cubic Å. Blue solid, red dashed and yellow dashed lines denote positive, negative and zero contours, respectively. **d**, Schematic representation of  $\text{Fe}_4\text{S}_4(\text{Cys-S}_\gamma)_4$ . Left, the plane of FE1–S4–FE2–S3; right, the plane of FE3–S1–FE4–S2. The atomic charges for each atom are indicated. The dashed lines indicate interactions between valence densities of sulfur atoms and hydrogen atoms, and the thin lines indicate interactions between valence densities of sulfur atoms and non-hydrogen atoms.

On the basis of the Atoms in Molecules (AIM) theory<sup>23</sup>, the bond topological properties can be derived from the charge-density analysis<sup>24</sup>. The Laplacian and gradient maps display bond paths as well as boundaries of the respective atoms. The crossing point of the bond path and the boundary between two atoms is defined as the bond critical point (BCP). Distortions of the bond paths were clearly observed in the Fe–S bonds of  $\text{Fe}_4\text{S}_4(\text{Cys-S}_\gamma)_4$  (Fig. 3a, b). The BCPs exist between all Fe–S bonds (Extended Data Table 5). The electron density at the BCP ( $\rho_{\text{BCP}}$ ) values for most Fe–S bonds have a negative correlation with bond lengths, as expected<sup>24,25</sup>. However, those for FE1–S2 and FE2–S1 do not follow this rule (Fig. 3c). These two bonds have low  $\rho_{\text{BCP}}$  values despite short bond lengths. This fact is consistent with weak 3*d*–3*p* overlaps as shown in the deformation map (Fig. 2). The bond strengths are unequal among the four Fe–S bonds of each Fe atom. The Fe–(Cys-S $\gamma$ ) bonds give the highest  $\rho_{\text{BCP}}$  values (Extended Data Table 5). In addition, the BCPs are located at closer positions to Fe atoms than the other three Fe–S bonds. This feature of Fe–(Cys-S $\gamma$ ) may give some contribution to the distortion of the  $\text{Fe}_4\text{S}_4$  cluster from the  $T_d$  symmetry, in addition to the spin coupling scheme<sup>26</sup> and the difference in bond length<sup>27</sup>. Cys46, which is the only cysteine residue showing the planar peptide bond, gives the lowest  $\rho_{\text{BCP}}$  among the four Fe–(Cys-S $\gamma$ ) bonds (Extended Data Table 5). This may be further evidence for the weaker interaction between FE2 and Cys46 as stated above. The Laplacian of electron densities ( $\nabla^2\rho$ ) at all BCPs in  $\text{Fe}_4\text{S}_4(\text{Cys-S}_\gamma)_4$  is positive, as in the case of other transition metal complexes<sup>22</sup>.

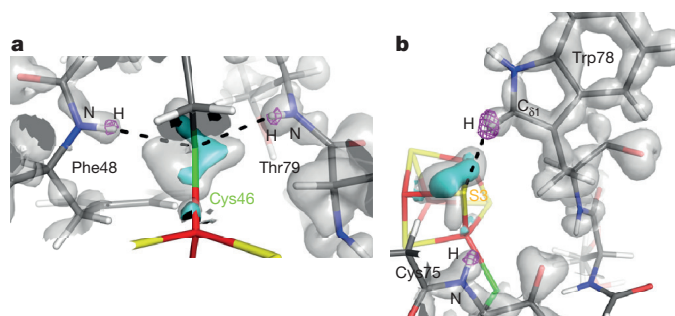
The contributions of valence electrons in the interactions between  $\text{Fe}_4\text{S}_4(\text{Cys-S}_\gamma)_4$  and the protein environment are clearly visualized in the charge-density analysis (Fig. 2d). Valence electrons of Cys-S $\gamma$  atoms





**Figure 3 | Topological analysis of charge density in the  $\text{Fe}_4\text{S}_4$  cluster.** **a**, The Laplacian  $\nabla^2\rho$  map in the plane of FE1–S3–FE4–S2. Contour interval is 0.05 electrons per ångström to the fifth power. Red solid lines represent negative values and blue dashed lines represent positive values. BCPs are shown as crosses. **b**, The gradient vector fields in the same plane as in **a**. **c**, Relationship between bond length and  $\rho_{\text{BCP}}$ . Filled circles in blue, pink and green are for intra-subcluster Fe–S, inter-subcluster Fe–S and Fe–(Cys–S<sub>γ</sub>), respectively.

of Cys43, Cys46, Cys61 and Cys75 interact with the carbonyl oxygen atom of Asn70 and H atom of Ile69–C<sub>δ1</sub> (Extended Data Fig. 3a), the main-chain amides of Phe48 and Thr79 (Fig. 4a), the main-chain amide of Leu63 and the H atom of Phe64–C<sub>δ2</sub> (Extended Data Fig. 3b), and the main-chain amide of Ser77 (Extended Data Fig. 3c), respectively. Valence electrons of S1, S2 and S3 interact with the H atom of Phe48–C<sub>δ2</sub> (Extended Data Fig. 3d), the H atoms of Phe64–C<sub>ε2</sub> and Ile69–C<sub>γ2</sub> (Extended Data Fig. 3e), and the H atom of Trp78–C<sub>δ1</sub> (Fig. 4b). No interactions are observed between S2 and the H atom of Tyr19–C<sub>δ1</sub> (Extended Data Fig. 3e), and S3 and the main-chain amide of Cys75 (Fig. 4b), although these are in close proximity. The S4 atom does not interact with the H atoms of Cys43–C<sub>β</sub> and Cys46–C<sub>β</sub> (Extended Data Fig. 3f). In addition, the S4 atom is not surrounded by aromatic side chains. The interaction between valence electrons and H atoms has correlation with atomic charges of S atoms. S4 with the largest atomic charge (–1.6) has no such interactions, while other bridging S atoms with smaller charges (–0.5 to –0.4) have (Fig. 2d). Furthermore, this rule is consistent for Cys–S atoms. Cys46–S<sub>γ</sub> and Cys61–S<sub>γ</sub> with smaller charges (–0.1 and –0.3) have two interactions, while Cys43–S<sub>γ</sub> and Cys75–S<sub>γ</sub> with larger charges (–1.5 and –1.3) have only one interaction. Charge transfer from S to H atoms through the interactions may be involved in the reduction of the negative charges of S. Taken together with atomic charges and topological properties, detailed views of interactions between the cluster and surrounding ligands imply that the FE1 atom together with FE2, S4 and Cys43–S<sub>γ</sub> atoms are crucial for storing electronic charges in the reduced HiPIP. The atomic charges in this paper were derived from the AIM analysis. It has been shown that the AIM analysis provides the most reliable estimation of the electronic properties for model clusters<sup>26</sup>. In our study, the absolute values of the AIM charges are generally larger than charges derived from the multipolar parameters (Extended Data Table 4). This can be attributed



**Figure 4 | Interaction network around the  $\text{Fe}_4\text{S}_4$  cluster.** **a**, Deformation electron density around the Cys46–S<sub>γ</sub> atom. The main-chain amides of Phe48 and Thr79 are located close to the Cys46–S<sub>γ</sub> atom. The static deformation maps are shown as grey and cyan surfaces contoured at the levels of +0.1 and +0.3 electrons per cubic ångström, respectively, and the omit map of hydrogen atoms is shown as a pink mesh contoured at the 3.0 $\sigma$  level. The dashed lines indicate interactions between valence densities of sulfur atoms and hydrogen atoms. **b**, Deformation electron density around S3 of the  $\text{Fe}_4\text{S}_4$  cluster. The main-chain amide of Cys75 and the H atom of Trp78–C<sub>δ1</sub> are located close to the S3 atom.

to charge transfer from S to Fe atoms. However, it is difficult to discuss this only from our results.

The cuboid  $[\text{Fe}_4\text{S}_4]^{2+}$  cluster is divided into two rhombic  $[\text{Fe}_2\text{S}_2]^{+}$  subclusters<sup>27</sup>. Each subcluster has ferromagnetically coupled two Fe atoms with  $d^9$  configuration ( $S = 9/2$ ). The two subclusters are anti-ferromagnetically coupled to each other to give  $S = 0$ . In HiPIP from *T. tepidum*, one subcluster consists of FE1, FE2, S3 and S4 atoms and coordinates Cys43 and Cys46, while another subcluster consists of FE3, FE4, S1 and S2 atoms and coordinates Cys61 and Cys75 by taking the results of HiPIP from *Chromatium vinosum*<sup>27,28</sup> into consideration. The assignment is consistent with the overlap manner in the Fe–S bonds and the topological analysis in which FE1–S2 and FE2–S1 between two subclusters show a distinctive feature. It was suggested that the redox reaction is localized in one of the two subclusters<sup>27</sup>. This is likely to be consistent with our implication for electron-storing atoms.

The charge-density analysis of this study experimentally reveals a subatomic structure of respective atoms in the polynuclear metal cluster in the HiPIP. This result, in combination with spectroscopic and computational data, will contribute to the understanding of the relationship between structure and function of metalloproteins.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 2 October 2015; accepted 12 April 2016.**

**Published online 18 May 2016.**

- Hendrickson, W. A. Stereochemically restrained refinement of macromolecular structures. *Methods Enzymol.* **115**, 252–270 (1985).
- Wlodawer, A., Minor, W., Dauter, Z. & Jaskolski, M. Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination. *FEBS J.* **280**, 5705–5736 (2013).
- Rokob, T. A., Srnc, M. & Rulíšek, L. Theoretical calculations of physico-chemical and spectroscopic properties of bioinorganic systems: current limits and perspectives. *Dalton Trans.* **41**, 5754–5768 (2012).
- Nogi, T., Fathir, I., Kobayashi, M., Nozawa, T. & Miki, K. Crystal structures of photosynthetic reaction center and high-potential iron-sulfur protein from *Thermochromatium tepidum*: thermostability and electron transfer. *Proc. Natl Acad. Sci. USA* **97**, 13561–13566 (2000).
- Liu, L., Nogi, T., Kobayashi, M., Nozawa, T. & Miki, K. Ultrahigh-resolution structure of high-potential iron-sulfur protein from *Thermochromatium tepidum*. *Acta Crystallogr. D* **58**, 1085–1091 (2002).
- Niwa, S. *et al.* Structure of the LH1–RC complex from *Thermochromatium tepidum* at 3.0 Å. *Nature* **508**, 228–232 (2014).
- Takeda, K., Kusumoto, K., Hirano, Y. & Miki, K. Detailed assessment of X-ray induced structural perturbation in a crystalline state protein. *J. Struct. Biol.* **169**, 135–144 (2010).
- Niu, S. & Ichiye, T. Insight into environmental effects on bonding and redox properties of  $[\text{4Fe-4S}]$  clusters in proteins. *J. Am. Chem. Soc.* **131**, 5724–5725 (2009).



9. Glaser, T. *et al.* Protein effects on the electronic structure of the  $[\text{Fe}_4\text{S}_4]^{2+}$  cluster in ferredoxin and HiPIP. *J. Am. Chem. Soc.* **123**, 4859–4860 (2001).
10. Dey, A. *et al.* Solvent tuning of electrochemical potentials in the active sites of HiPIP versus ferredoxin. *Science* **318**, 1464–1468 (2007).
11. Jelsch, C. *et al.* Accurate protein crystallography at ultra-high resolution: valence electron distribution in crambin. *Proc. Natl Acad. Sci. USA* **97**, 3171–3176 (2000).
12. Schmidt, A., Jelsch, C., Ostergaard, P., Rypniewski, W. & Lamzin, V. S. Trypsin revisited: crystallography at (sub) atomic resolution and quantum chemistry revealing details of catalysis. *J. Biol. Chem.* **278**, 43357–43362 (2003).
13. Fournier, B. *et al.* Charge density and electrostatic interactions of fidarestat, an inhibitor of human aldose reductase. *J. Am. Chem. Soc.* **131**, 10929–10941 (2009).
14. Zarychta, B. *et al.* Cholesterol oxidase: ultrahigh-resolution crystal structure and multipolar atom model-based analysis. *Acta Crystallogr. D* **71**, 954–968 (2015).
15. Hansen, N. K. & Coppens, P. Testing aspherical atom refinements on small-molecule data sets. *Acta Crystallogr. A* **34**, 909–921 (1978).
16. Berkholz, D. S., Driggers, C. M., Shapovalov, M. V., Dunbrack, R. L., Jr & Karplus, P. A. Nonplanar peptide bonds in proteins are common and conserved but not biased toward active sites. *Proc. Natl Acad. Sci. USA* **109**, 449–453 (2012).
17. Engh, R. A. & Huber, R. Accurate bond and angle parameters for X-ray structure refinement. *Acta Crystallogr. A* **47**, 392–400 (1991).
18. Murray-Rust, P. & Glusker, J. P. Directional hydrogen bonding to  $\text{sp}^2$ - and  $\text{sp}^3$ - hybridized oxygen atoms and its relevance to ligand-macromolecule interactions. *J. Am. Chem. Soc.* **106**, 1018–1025 (1984).
19. Bertini, I., Donaire, A., Felli, I. C., Luchinat, C. & Rosato, A.  $^1\text{H}$  and  $^{13}\text{C}$  NMR studies of an oxidized HiPIP. *Inorg. Chem.* **36**, 4798–4803 (1997).
20. Improta, R., Vitagliano, L. & Esposito, L. Peptide bond distortions from planarity: new insights from quantum mechanical calculations and peptide/protein crystal structures. *PLoS ONE* **6**, e24533 (2011).
21. Wang, Y.-F., Yu, Z.-Y., Wu, J. & Liu, C.-B. Electron delocalization and charge transfer in polypeptide chains. *J. Phys. Chem. A* **113**, 10521–10526 (2009).
22. Smith, G. T. *et al.* Experimental determination of the electron density topology in a non-centrosymmetric transition metal complex:  $[\text{Ni}(\text{H}_3\text{L})][\text{NO}_3][\text{PF}_6]$  [ $\text{H}_3\text{L} = N,N',N''$ -tris(2-hydroxy-3-methylbutyl)-1,4,7-triazacyclononane]. *J. Am. Chem. Soc.* **119**, 5028–5034 (1997).
23. Bader, R. F. W. *Atoms in Molecules: A Quantum Theory* (Oxford Univ. Press, 1990).
24. Koritsanszky, T. S. & Coppens, P. Chemical applications of X-ray charge-density analysis. *Chem. Rev.* **101**, 1583–1628 (2001).
25. Gibbs, G. V. *et al.* Theoretical electron density distributions for Fe- and Cu-sulfide earth materials: a connection between bond length, bond critical point properties, local energy densities, and bonded interactions. *J. Phys. Chem. B* **111**, 1923–1931 (2007).
26. Harris, T. V. & Szilagy, R. K. Iron-sulfur bond covalency from electronic structure calculations for classical iron-sulfur clusters. *J. Comput. Chem.* **35**, 540–552 (2014).
27. Dey, A. *et al.* Sulfur K-edge XAS and DFT calculations on  $[\text{Fe}_4\text{S}_4]^{2+}$  clusters: effects of H-bonding and structural distortion on covalency and spin topology. *Inorg. Chem.* **44**, 8349–8354 (2005).
28. Mouesca, J. M., Noodleman, L., Case, D. A. & Lamotte, B. Spin densities and spin coupling in iron-sulfur clusters: a new analysis of hyperfine coupling constants. *Inorg. Chem.* **34**, 4347–4359 (1995).

**Acknowledgements** We thank K. Kusumoto and H. Ohno for their contributions in the initial steps of the work, and T. Tsujinaka and S. Niwa for their contributions in the preparation of the manuscript. We also thank the BL41XU beamline staff of SPring-8 for their help in data collection. This work was supported by a Grant-in-Aid for Scientific Research (number 23657073 to K.T.) and the Photon and Quantum Basic Research Coordinated Development Program (to K.M.) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

**Author Contributions** K.M. initiated and supervised the project. K.T. designed the experiments. Y.H. prepared crystals. Y.H. and K.T. performed data collection and the crystallographic analysis. Y.H., K.T. and K.M. discussed the results. Y.H. wrote the initial draft, and K.T. and K.M. revised the manuscript.

**Author Information** The coordinates and structure factors have been deposited in the Protein Data Bank under accession number 5D8V. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to K.M. ([miki@kuchem.kyoto-u.ac.jp](mailto:miki@kuchem.kyoto-u.ac.jp)).

## METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

**X-ray diffraction experiment.** Crystals of DTT-reduced HiPIP were prepared as previously described<sup>7</sup>. The X-ray wavelengths were set to 0.45 Å (27.6 keV) at the BL41XU beamline of SPring-8 (proposal numbers 2008B-1337, 2010A1237 and 2010B1284 to K.T.). The diffraction intensities were measured using a Rayonix MX-225 CCD detector. The crystals were cooled during the data collection at 100 K using a nitrogen-gas stream. Three data sets for reflections in ultra-high, medium- and low-resolution regions were separately collected for the different positions of a single crystal (0.8 mm × 0.2 mm × 0.1 mm), and the data sets for reflections in medium and low-resolution regions were collected with attenuated X-ray exposure. The data set for ultra-high resolution was collected using a helical data collection procedure<sup>29</sup> with a microbeam of 50 μm × 50 μm. The maximum dose in each irradiated position was limited to ~6 × 10<sup>5</sup> Gy. The dose was estimated using the RADDOSE program<sup>30</sup>. The diffraction data sets were integrated and scaled using the HKL2000 program package<sup>31</sup>. The relative *B* factors, *R*<sub>sym</sub> and other crystallographic statistics, which represent the metrics for radiation damage, did not show significant changes during the data collection (Extended Data Fig. 1). Three data sets for ultra-high-, medium-, and low-resolution regions were merged into a complete data set.

**Structure refinement with the ISAM.** The test set for free *R* factor calculation was set to be the same as in the 0.7 Å-resolution data of HiPIP<sup>7</sup> in the resolution range 20–0.7 Å and extended to 0.48 Å resolution by randomly selecting 5% of reflections. Structure refinement was started from the low-dose structure of the reduced form of HiPIP (Protein Data Bank accession number 3A39) using the SHELXL program<sup>32</sup>. Positional and anisotropic displacement parameters were refined for all non-hydrogen atoms. Restraints for bond lengths and angles were removed for non-hydrogen atoms of single conformational residues and for the Fe<sub>4</sub>S<sub>4</sub> cluster. All of the hydrogen atoms of amino-acid residues were included in the model, and only the hydrogen atoms of water molecules which were observed in the *F*<sub>o</sub> – *F*<sub>c</sub> map contoured at the 2.0σ level were included. Bond lengths and angles involving hydrogen atoms were constrained using the riding model in the SHELXL program. The final structure in the SHELXL refinement contains 34 multiple-conformation residues and was refined to the *R*<sub>work</sub> and *R*<sub>free</sub> factors of 8.24% and 8.63%.

**Charge-density analysis with the MAM.** The charge-density analysis with the MAM was performed using the MoPro program<sup>33</sup>. The refinement for HiPIP followed the procedure for human aldose reductase<sup>34</sup>, but was slightly modified. The scale factors were refined in the whole resolution range 20–0.48 Å. Bulk solvent parameters with the scale factors were refined in the resolution range 20–1.0 Å. The subsequent refinement processes were performed for the non-hydrogen atoms of single conformational residues, the atoms in the Fe<sub>4</sub>S<sub>4</sub> cluster and the oxygen atoms of water molecules with two hydrogen atoms (H<sub>2</sub>O). Positional and anisotropic displacement parameters were refined in the whole resolution range. Then, positional and anisotropic displacement parameters were refined using only high-resolution reflections (high-order refinement). In the high-order refinement, resolution ranges used for refinement were sequentially changed as 1.0–0.48 Å, 0.9–0.48 Å, 0.8–0.48 Å, 0.7–0.48 Å and 0.65–0.48 Å. Positions of hydrogen atoms were changed to the standard geometry determined by neutron diffraction experiments of small molecules<sup>35</sup> and were fixed in the subsequent refinement steps. After the high-order refinement, strong peaks remained near the atom positions of the Fe<sub>4</sub>S<sub>4</sub> cluster in the residual electron-density map. The short path lengths of low-resolution reflections cause an incomplete absorption of the high-energy X-rays in the CCD detector<sup>36</sup>. To correct the resolution-dependence in the absorption of the CCD detector, we used an exponential function by which the observed structure factors were fitted to the calculated ones as  $F_{\text{obs}}^{\text{corr}}(hkl) = F_{\text{obs}}(hkl)/EF$  where  $EF = \sum_i a_i \exp[b_i(\sin\theta/\lambda - c_i)] + d$ , and the coefficients of *a<sub>i</sub>*, *b<sub>i</sub>*, *c<sub>i</sub>* (*i* = 1, 2, 3) and *d* were fitted to <|*F*<sub>o</sub>|/|*F*<sub>c</sub>> in 50 resolution bins by a least-squares method.

The MAM is expressed as shown in equation (1):

$$\rho_{\text{atom}}(\mathbf{r}) = \rho_{\text{core}}(r) + P_{\text{val}}\kappa^3\rho_{\text{val}}(\kappa r) + \sum_{l=0}^{l_{\text{max}}} \kappa'^3 R_{nl}(\kappa' r) \sum_{m=0}^l P_{lm\pm} y_{lm\pm}(\theta, \varphi) \quad (1)$$

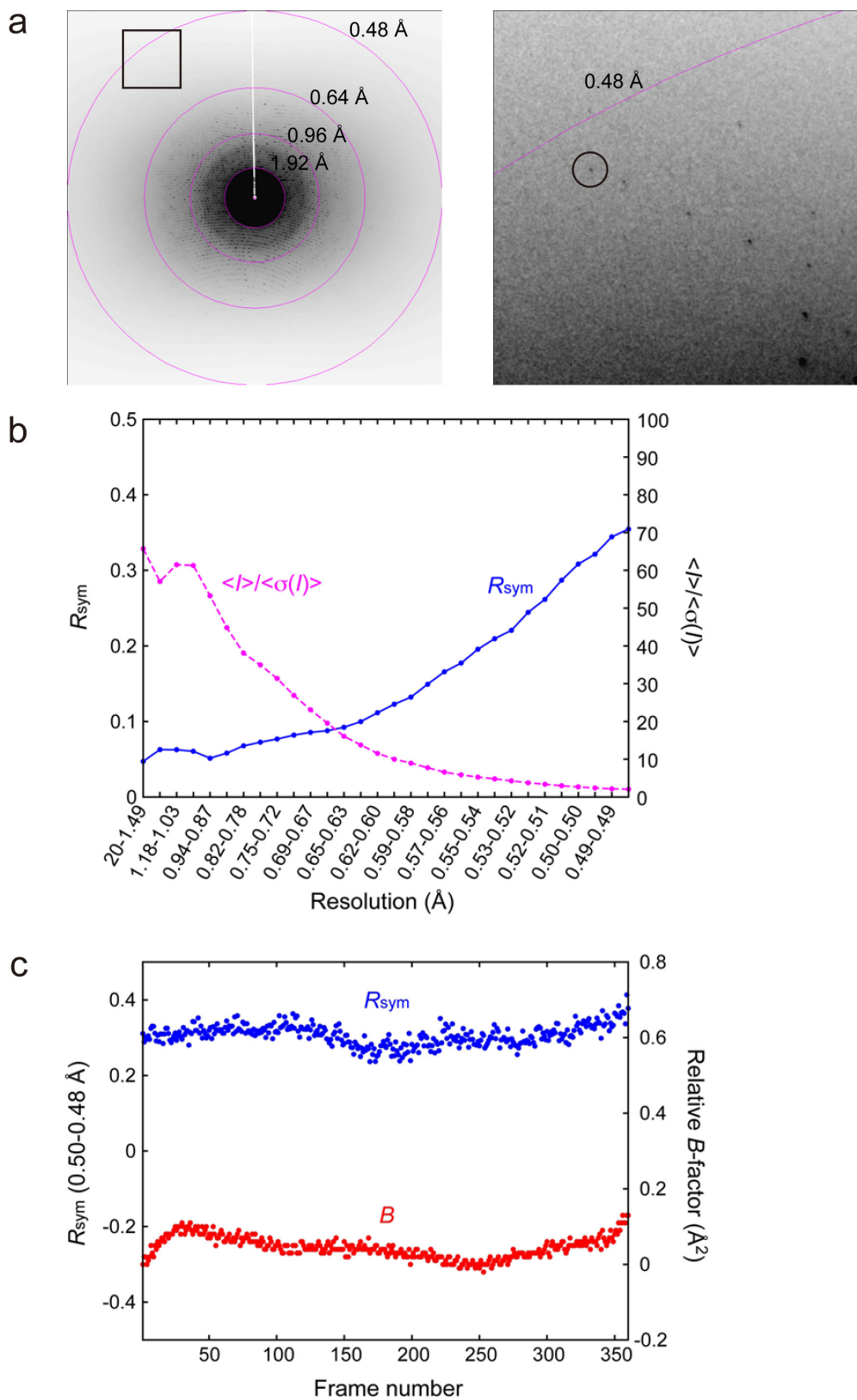
The  $\rho_{\text{atom}}$ ,  $\rho_{\text{core}}$  and  $\rho_{\text{val}}$  represent the total, spherical core and spherical valence electron densities, and  $P_{\text{val}}$  and  $P_{lm\pm}$  are the spherical valence and deformation multipole populations.  $R_{nl}$  is the Slater-type radial functions and  $y_{lm\pm}$  the real spherical harmonics. The  $\kappa$  and  $\kappa'$  parameters are electron-density expansion/contraction coefficients. The parameters used for the least-squares refinement are  $P_{\text{val}}$ ,  $P_{lm\pm}$ ,  $\kappa$  and  $\kappa'$ . Multipolar parameters were transferred to all of the amino-acid residues and H<sub>2</sub>O molecules from the experimental library multipolar atom model (ELMAM)<sup>37</sup>. The multipolar parameters were assigned to the octupole

level ( $l_{\text{max}} = 3$ ) for C, N and O atoms, the hexadecapole level ( $l_{\text{max}} = 4$ ) for S and Fe atoms and the dipole level ( $l_{\text{max}} = 1$ ) for H atoms. Solvent atoms other than H<sub>2</sub>O were treated as spherical with a neutral charge. The definition of local axes was derived from the multipolar library for the atoms of the amino-acid residues and H<sub>2</sub>O. The local axes for the sulfur atoms of Fe<sub>4</sub>S<sub>4</sub>(Cys-S-)<sub>4</sub> were defined to coincide with the directions of the unit cell axes. The *d*-orbital populations for the local axes (*x*, *y*, *z*) were calculated and are listed in Extended Data Table 6. The local axes were set as shown in Extended Data Fig. 4, according to refs 38, 39.

Initial values of multipolar parameters for Fe<sub>4</sub>S<sub>4</sub>(Cys-S-)<sub>4</sub> were unavailable in ELMAM. However, the atomic charge of the Fe atom is difficult to define properly in the multipolar refinement because of the diffused distribution of the 4s electrons<sup>38</sup>. In the multipolar refinement of HiPIP, the initial values of  $P_{\text{val}}$  and  $\kappa$  parameters were obtained from a grid search procedure for Fe, S and Cys-S- atoms. The  $\kappa'$  parameter was fixed to 1.00. All of the  $P_{lm\pm}$  parameters ( $l > 0$ ) were refined for the atoms of Fe<sub>4</sub>S<sub>4</sub>(Cys-S-)<sub>4</sub>, but the monopole  $P_{00}$  parameters were fixed to zero because the  $P_{00}$  parameter has a high correlation with the  $P_{\text{val}}$  parameter. Finally, the initial values of  $P_{\text{val}}$  and  $\kappa$  parameters were determined to be 7.5 and 1.10 for Fe atoms, 6.5 and 0.99 for S atoms and 6.5 and 0.99 for Cys-S- atoms.

After the grid search, the  $P_{lm\pm}$  parameters were refined for single conformational residues, Fe<sub>4</sub>S<sub>4</sub>(Cys-S-)<sub>4</sub> and H<sub>2</sub>O molecules. Following the refinement of the  $P_{lm\pm}$  parameters, positional and anisotropic displacement parameters were refined for the non-hydrogen atoms. In the refinement of the  $P_{\text{val}}$  parameters it was difficult to obtain reasonable values for the atomic charges in Fe<sub>4</sub>S<sub>4</sub>(Cys-S-)<sub>4</sub>. Thus, the  $P_{\text{val}}$  parameters were refined after the  $P_{lm\pm}$  parameters. The differences in the atomic scattering factors for Fe<sup>2+</sup> and Fe<sup>3+</sup> were prominent in the resolution range lower than 1.5 Å. Reasonable atomic charges were obtained in the refinement of  $P_{\text{val}}$  parameters with the resolution range 20–1.2 Å. The  $P_{\text{val}}$  parameters for single conformational residues were also refined, but the  $P_{\text{val}}$  parameters for hydrogen atoms were constrained to be the same as for the chemically equivalent atoms. The total number of electrons in the crystal was fixed during the refinement of  $P_{\text{val}}$  parameters. Finally, the *R*<sub>work</sub> and *R*<sub>free</sub> factors were converged to 7.16% and 7.80% (Extended Data Table 1). *B*<sub>eq</sub> values for Fe<sub>4</sub>S<sub>4</sub>(Cys-S-)<sub>4</sub> in the final model are listed in Extended Data Table 4. The static deformation map in this paper was calculated by the equation  $\Delta\rho_{\text{static}}(\mathbf{r}) = \sum_{j=1}^{N_{\text{at}}} [\rho_{\text{multi}}(\mathbf{r} - \mathbf{r}_j) - \rho_{\text{IAM}}(\mathbf{r} - \mathbf{r}_j)]$ . The parameters  $\rho_{\text{multi}}$  and  $\rho_{\text{IAM}}$  represent the electron density calculated from the MAM and from the ISAM, respectively. The static deformation map is calculated excluding the contribution of the atomic displacement parameters of atoms. The two-dimensional contour maps were prepared using the VMOPRO program and the three-dimensional figures were prepared using PyMOL<sup>40</sup>. Topological analyses based on the AIM theory<sup>23,24</sup> was performed with VMOPRO. The atomic charges by the AIM theory were calculated with Bader<sup>41</sup>. The atomic charges by the equation of  $q = N_{\text{val}} - P_{\text{val}}$ , where  $N_{\text{val}}$  is the number of valence electrons in the neutral charge, were also calculated. The two methods gave highly correlated results as listed in Extended Data Table 4.

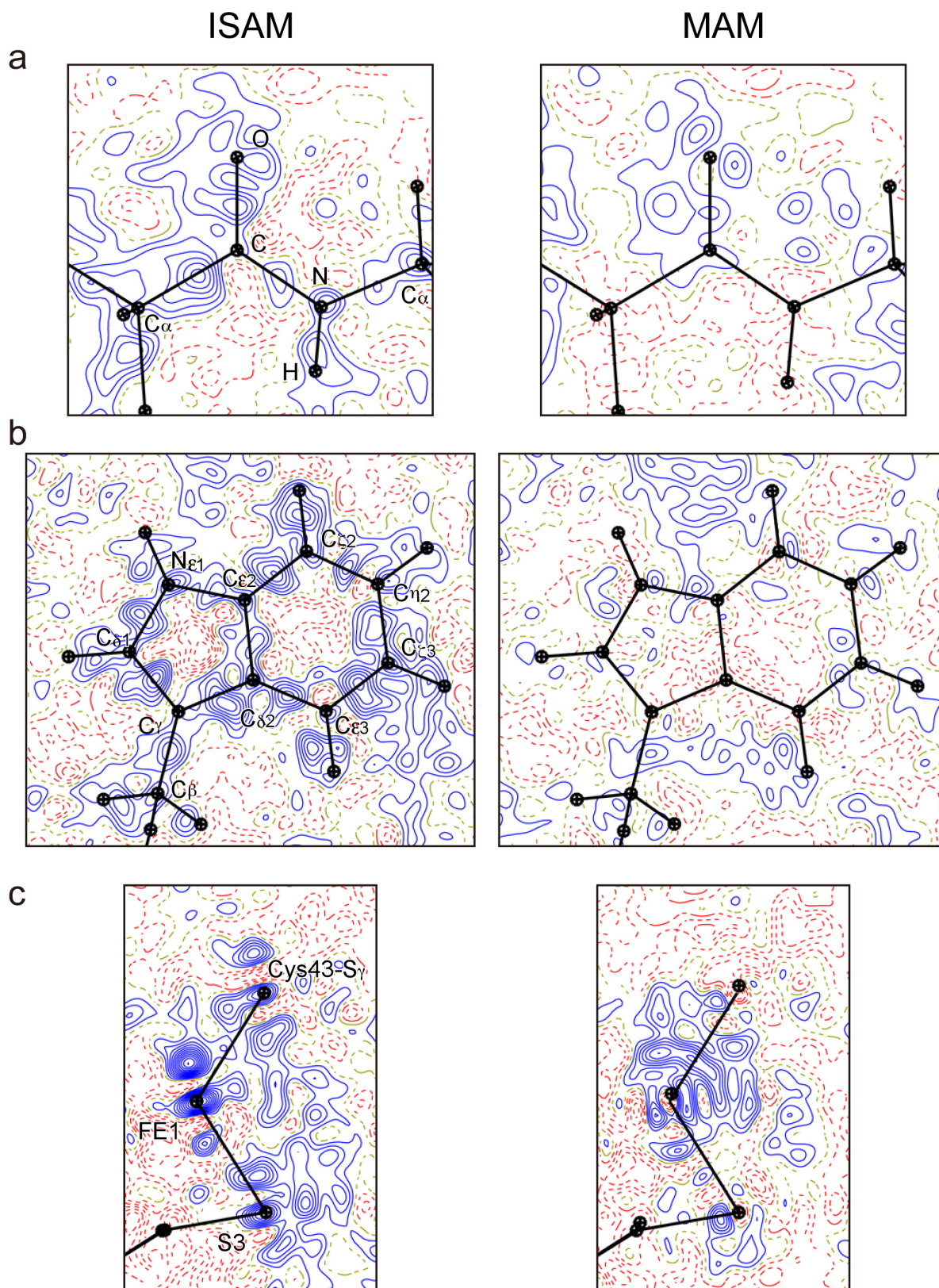
29. Flot, D. *et al.* The ID23-2 structural biology microfocus beamline at the ESRF. *J. Synchrotron Radiat.* **17**, 107–118 (2010).
30. Paithankar, K. S., Owen, R. L. & Garman, E. F. Absorbed dose calculations for macromolecular crystals: improvements to RADDOSE. *J. Synchrotron Radiat.* **16**, 152–162 (2009).
31. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data. *Methods Enzymol.* **276**, 307–326 (1997).
32. Sheldrick, G. M. A short history of SHELX. *Acta Crystallogr. A* **64**, 112–122 (2008).
33. Guillot, B., Viry, L., Guillot, R. & Lecomte, C. Refinement of proteins at subatomic resolution with MOPRO. *J. Appl. Cryst.* **34**, 214–223 (2001).
34. Guillot, B., Jelsch, C., Podjarny, A. & Lecomte, C. Charge-density analysis of a protein structure at subatomic resolution: the human aldose reductase case. *Acta Crystallogr. D* **64**, 567–588 (2008).
35. Allen, F. H. A systematic pairwise comparison of geometric parameters obtained by X-ray and neutron diffraction. *Acta Crystallogr. B* **42**, 515–522 (1986).
36. Wu, G., Rodrigues, B. L. & Coppens, P. The correction of reflection intensities for incomplete absorption of high-energy X-rays in the CCD phosphor. *J. Appl. Cryst.* **35**, 356–359 (2002).
37. Zarychta, B., Pichon-Pesme, V., Guillot, B., Lecomte, C. & Jelsch, C. On the application of an experimental multipolar pseudo-atom library for accurate refinement of small-molecule and protein crystal structures. *Acta Crystallogr. A* **63**, 108–125 (2007).
38. Holladay, A., Leung, P. & Coppens, P. Generalized relations between *d*-orbital occupancies of transition-metal atoms and electron-density multipole population parameters from X-ray diffraction data. *Acta Crystallogr. A* **39**, 377–387 (1983).
39. Sabino, J. R. & Coppens, P. On the choice of *d*-orbital coordinate system in charge-density studies of low-symmetry transition-metal complexes. *Acta Crystallogr. A* **59**, 127–131 (2003).
40. DeLano, W. L. The PyMol Molecular Graphics System (DeLano Scientific, 2002).
41. Yu, M. & Trinkle, D. R. Accurate and efficient algorithm for Bader charge integration. *J. Chem. Phys.* **134**, 064111 (2011).



**Extended Data Figure 1 | Quality of the diffraction data at 0.48 Å resolution.** **a**, The diffraction image. Right, zoom view of the boxed region at left. The resolution for each circle is indicated. **b**,  $R_{\text{sym}}$  (blue) and  $\langle I \rangle / \langle \sigma(I) \rangle$  (pink) values are plotted for 30 resolution bins. **c**, Changes

of  $R_{\text{sym}}$  at the highest-resolution shell (0.50–0.48 Å) and relative  $B$  factor in the course of the data collection. The  $R_{\text{sym}}$  (blue) and relative  $B$  factor (red) are plotted as functions of frame number.



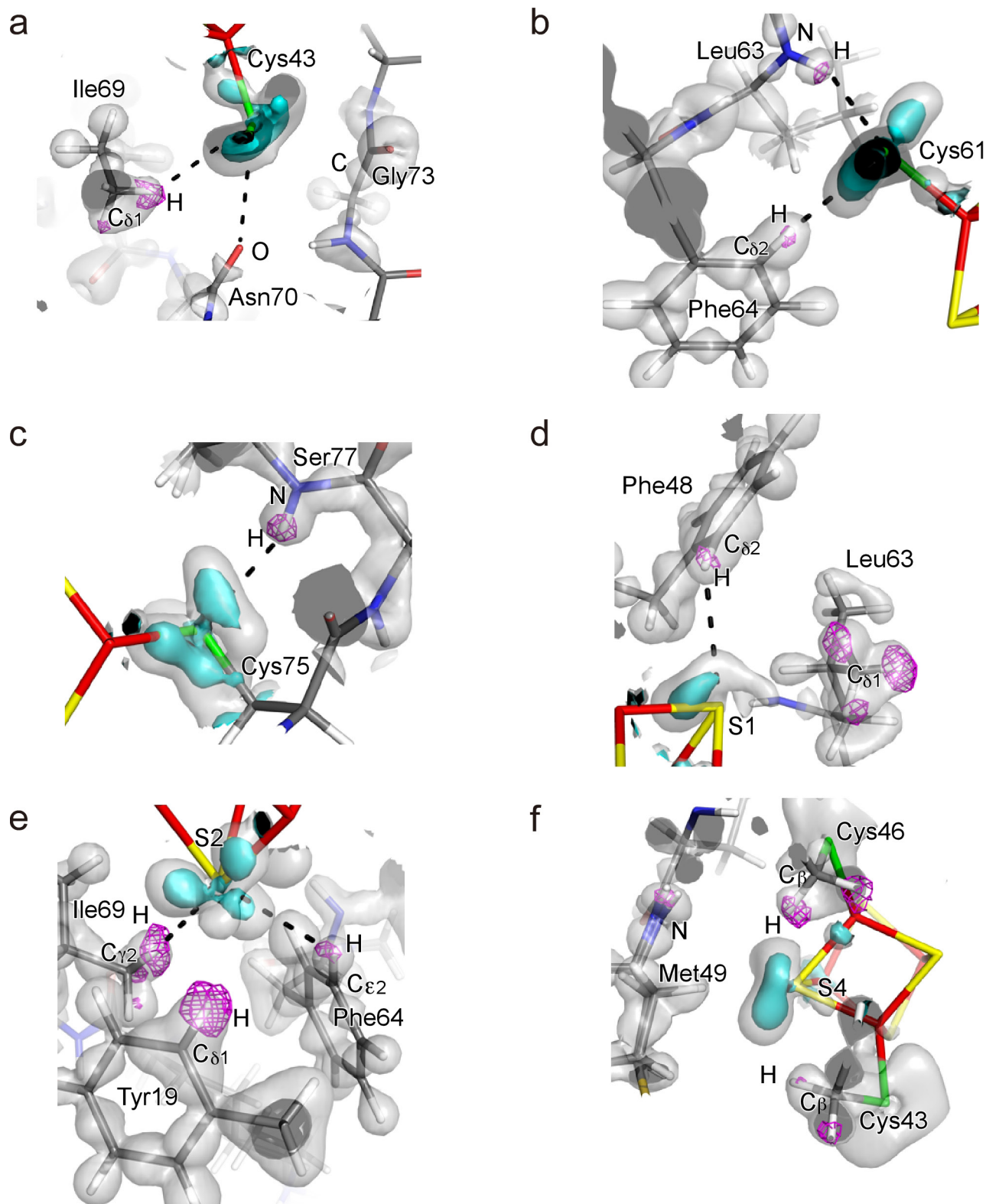


**Extended Data Figure 2 | Residual electron density for each refinement step.** The left panels show the residual density after the ISAM refinement; the right panels show the residual density after the MAM refinement.

**a,** The plane of the peptide bond between Asn45 and Cys46. Maximum and minimum peaks are 0.33 and  $-0.22$  electrons per cubic ångström for the ISAM analysis, and 0.18 and  $-0.20$  electrons per cubic ångström for the MAM analysis. **b,** The plane of the aromatic ring of Trp74. Maximum and minimum peaks are 0.34 and  $-0.29$  electrons per cubic ångström

for the ISAM analysis, and 0.23 and  $-0.23$  electrons per cubic ångström for the MAM analysis. **c,** The  $\text{Fe}_4\text{S}_4$  cluster. The plane consists of FE1, S3 and Cys43-S $\gamma$  atoms. Maximum and minimum peaks are 0.60 and  $-0.35$  electrons per cubic ångström for the ISAM analysis, and 0.35 and  $-0.29$  electrons per cubic ångström for the MAM analysis. The contour interval is 0.05 electrons per cubic ångström for all figures. Blue solid, red dashed and yellow dashed lines denote positive, negative and zero contours, respectively.

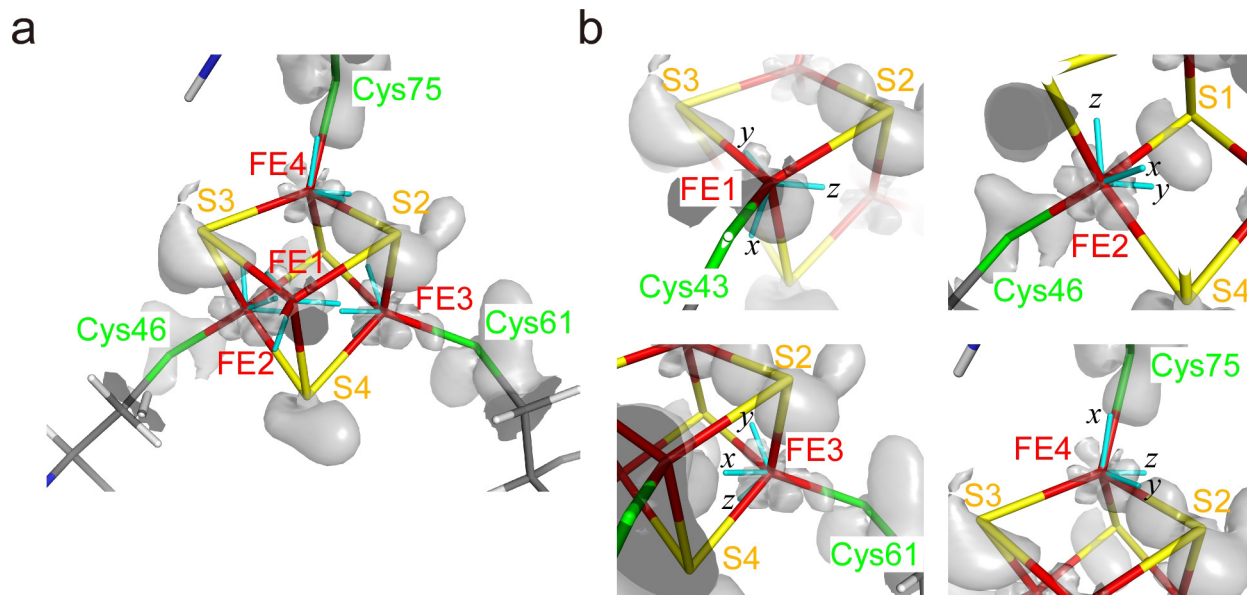




### Extended Data Figure 3 | Interaction network around the $\text{Fe}_4\text{S}_4$ cluster.

**a**, Deformation electron density around the Cys43- $\text{S}_\gamma$  atom. The main-chain oxygen atom of Asn70, the main-chain carboxyl carbon atom of Gly73 and the H atom of Ile69- $\text{C}_{\delta 1}$  are located close to Cys43- $\text{S}_\gamma$ . The static deformation maps are shown as grey and cyan surfaces contoured at the levels of +0.1 and +0.3 electrons per cubic ångström, respectively. The omit map of hydrogen atoms is shown as a pink mesh contoured at the  $3.0\sigma$  level. The dashed lines indicate interactions between valence densities of sulfur atoms and hydrogen atoms. **b**, Deformation electron density around Cys61- $\text{S}_\gamma$ . The main-chain amide of Leu63 and the H atom of

Phe64- $\text{C}_{\delta 2}$  are located close to Cys61- $\text{S}_\gamma$ . **c**, Deformation electron density around Cys75- $\text{S}_\gamma$ . The main-chain amide of Ser77 is located close to Cys75- $\text{S}_\gamma$ . **d**, Deformation electron density around S1 of the  $\text{Fe}_4\text{S}_4$  cluster. The H atom of Phe48- $\text{C}_{\delta 2}$  and the  $\text{C}_{\delta 1}$  atom of Leu63 are located close to S1. **e**, Deformation electron density around S2 of the  $\text{Fe}_4\text{S}_4$  cluster. The H atoms of Tyr19- $\text{C}_{\delta 1}$ , Phe64- $\text{C}_{\epsilon 2}$  and Ile69- $\text{C}_{\gamma 2}$  are located close to S2. **f**, Deformation electron density around S4 of the  $\text{Fe}_4\text{S}_4$  cluster. The H atom of Cys43- $\text{C}_\beta$ , the H atom of Cys46- $\text{C}_\beta$  and the amide nitrogen atom of Met49 are located close to S4.



**Extended Data Figure 4 | The local axes for Fe atoms of  $\text{Fe}_4\text{S}_4(\text{Cys-S}_\gamma)_4$ .** **a**, Whole view of the local axes of the four Fe atoms. **b**, Close-up views of the local axes of each Fe atom (FE1–FE4). The static deformation maps of  $\text{Fe}_4\text{S}_4(\text{Cys-S}_\gamma)_4$  are represented as grey isosurfaces contoured at the level of +0.2 electrons per cubic ångström.

Extended Data Table 1 | Data collection and refinement statistics

	HiPIP
<b>Data collection</b>	
Space group	$P2_12_12_1$
Cell dimensions	
$a, b, c$ (Å)	46.48, 58.91, 23.44
$\alpha, \beta, \gamma$ (°)	90, 90, 90
Resolution (Å)	20–0.48 (0.50–0.48)*
$R_{\text{sym}}$ or $R_{\text{merge}}$ (%)	5.6 (33.9)
$I/\sigma I$	61.1 (2.7)
Completeness (%)	96.3 (89.0)
Redundancy	5.4 (3.0)
<b>Refinement</b>	
Resolution (Å)	20–0.48
No. reflections	301,119
$R_{\text{work}}/R_{\text{free}}$ (%) (ISAM)	8.24/8.63
$R_{\text{work}}/R_{\text{free}}$ (%) (MAM)	7.16/7.80
No. non-H atoms	
Protein	848
Ligand/ion	72
Water	185
No. H atoms	
Protein	820
Ligand/ion	0
Water	42

\*Highest-resolution shell is shown in parentheses.

Extended Data Table 2 | Dihedral and improper angles

	Dihedral angle (°)	Improper angle* (°)
Tyr19–Asn20	162.0	1.78
Asp22–Ala23	-177.0	0.98
Ala23–Thr24	177.0	1.01
Ala31–Ala32	176.4	1.20
Gln41–His42	168.3	0.01
His42–Cys43	-170.9	2.34
Cys43–Ala44	-175.2	2.52
Ala44–Asn45	-179.5	0.96
Gly60–Cys61	166.7	1.08
Leu68–Ile69	170.0	1.77
Gly73–Thr74	179.6	0.66
Thr74–Cys75	163.7	0.01
Cys75–Ala76	-172.8	2.18
Ala76–Ser77	170.7	1.10

\*The improper angle is an angle between the C–N'–C<sub>α</sub>' and H'–N'–C<sub>α</sub>' planes.



Extended Data Table 3 | Geometrical parameters in  $\text{Fe}_4\text{S}_4(\text{Cys-S}_\gamma)_4$ 

Distance (Å)		Angle (°)	
FE1-S2	2.2256 (3)*	FE1-S2-FE4	73.12 (1)*
FE1-S4	2.3096 (4)	FE1-S2-FE3	73.53 (-)
FE1-S3	2.3212 (4)	FE1-S4-FE3	73.13 (1)
FE2-S1	2.2177 (3)	FE1-S4-FE2	72.81 (1)
FE2-S3	2.3064 (3)	FE1-S3-FE4	71.94 (1)
FE2-S4	2.3185 (4)	FE1-S3-FE2	72.81 (1)
FE3-S4	2.2489 (3)	FE2-S1-FE4	73.68 (-)
FE3-S2	2.3110 (4)	FE2-S1-FE3	73.14 (2)
FE3-S1	2.3150 (4)	FE2-S3-FE4	72.54 (2)
FE4-S3	2.2705 (3)	FE2-S4-FE3	72.51 (1)
FE4-S1	2.2972 (4)	FE3-S2-FE4	72.31 (1)
FE4-S2	2.3015 (4)	FE3-S1-FE4	72.31 (1)
FE1-FE2	2.7466 (3)	S1-FE2-S3	104.73 (2)
FE1-FE3	2.7161 (2)	S1-FE2-S4	105.32 (1)
FE1-FE4	2.6974 (3)	S1-FE4-S3	103.34 (1)
FE2-FE3	2.7018 (3)	S1-FE4-S2	106.39 (1)
FE2-FE4	2.7077 (3)	S1-FE3-S4	104.42 (1)
FE3-FE4	2.7212 (3)	S1-FE3-S2	105.49 (1)
FE1-(Cys43-S <sub>γ</sub> )	2.2563 (3)	S2-FE1-S4	104.35 (2)
FE2-(Cys46-S <sub>γ</sub> )	2.2824 (3)	S2-FE1-S3	105.51 (1)
FE3-(Cys61-S <sub>γ</sub> )	2.2626 (4)	S2-FE4-S3	104.70 (1)
FE4-(Cys75-S <sub>γ</sub> )	2.2709 (3)	S2-FE3-S4	103.56 (1)
(Tyr19-C <sub>δ1</sub> )-S2	3.673 (2)	S3-FE2-S4	105.06 (1)
(Phe48-N)-(Cys46-S <sub>γ</sub> )	3.444 (1)	S3-FE1-S4	104.87 (1)
(Phe48-C <sub>δ2</sub> )-S1	3.792 (2)	(Cys43-S <sub>γ</sub> )-FE1-S2	113.49 (1)
(Leu63-N)-(Cys61-S <sub>γ</sub> )	3.375 (1)	(Cys43-S <sub>γ</sub> )-FE1-S4	111.60 (2)
(Leu63-C <sub>δ1</sub> )-S1	3.525 (2)	(Cys43-S <sub>γ</sub> )-FE1-S3	115.99 (2)
(Phe64-C <sub>δ2</sub> )-(Cys61-S <sub>γ</sub> )	3.716 (2)	(Cys46-S <sub>γ</sub> )-FE2-S1	114.63 (1)
(Phe64-C <sub>ε2</sub> )-S2	3.964 (2)	(Cys46-S <sub>γ</sub> )-FE2-S3	115.38 (1)
(Ile69-C <sub>γ2</sub> )-S2	3.893 (2)	(Cys46-S <sub>γ</sub> )-FE2-S4	110.79 (1)
(Asn70-O)-(Cys43-S <sub>γ</sub> )	3.357 (1)	(Cys61-S <sub>γ</sub> )-FE3-S4	117.52 (2)
(Cys75-N)-S3	3.401 (1)	(Cys61-S <sub>γ</sub> )-FE3-S2	119.16 (2)
(Ser77-N)-(Cys75-S <sub>γ</sub> )	3.365 (1)	(Cys61-S <sub>γ</sub> )-FE3-S1	105.22 (2)
(Trp78-C <sub>δ1</sub> )-S3	3.771 (1)	(Cys75-S <sub>γ</sub> )-FE4-S3	126.49 (2)
(Thr79-N)-(Cys46-S <sub>γ</sub> )	3.526 (1)	(Cys75-S <sub>γ</sub> )-FE4-S1	104.97 (2)
		(Cys75-S <sub>γ</sub> )-FE4-S2	109.41 (1)

\*Values in parentheses are estimated standard deviations given by full-matrix least-squares refinement.

**Extended Data Table 4 | Atomic properties of  $\text{Fe}_4\text{S}_4(\text{Cys-S}_\gamma)_4$** 

Atom	Charge (AIM)	Charge ( $N-P_{\text{val}}$ )	$B_{\text{eq}}$ factor ( $\text{\AA}^2$ )
FE1	+0.92	+0.42 (12)*	2.32
FE2	+1.15	+0.58 (12)	2.31
FE3	+1.13	+0.97 (12)	2.40
FE4	+1.46	+1.33 (12)	2.33
S1	-0.39	-0.14 (27)	2.53
S2	-0.51	-0.36 (27)	2.47
S3	-0.52	-0.21 (28)	2.40
S4	-1.57	-1.40 (28)	2.46
Cys43-S $_\gamma$	-1.48	-1.38 (28)	2.51
Cys46-S $_\gamma$	-0.12	-0.16 (28)	2.36
Cys61-S $_\gamma$	-0.27	-0.20 (28)	2.82
Cys75-S $_\gamma$	-1.27	-1.11 (29)	2.44
Total	-1.47	-1.69	

\*Values in parentheses are estimated standard deviations given by full-matrix least-squares refinement.

Extended Data Table 5 | Topological parameters at BCPs of Fe–S bonds

	$d$ (Å)	$d_1^*$ (Å)	$\rho$ (e/Å <sup>3</sup> )	$\nabla^2\rho$ (e/Å <sup>5</sup> )
FE1-S2	2.2256	1.058	0.54	1.4
FE1-S3	2.3212	1.024	0.61	3.7
FE1-S4	2.3096	1.093	0.44	3.4
FE2-S1	2.2177	0.966	0.61	5.2
FE2-S3	2.3064	1.056	0.51	2.5
FE2-S4	2.3185	1.131	0.50	2.2
FE3-S1	2.3150	1.050	0.65	2.3
FE3-S2	2.3110	1.049	0.59	1.7
FE3-S4	2.2489	0.973	0.73	5.5
FE4-S1	2.2972	1.056	0.54	2.0
FE4-S2	2.3015	1.024	0.70	3.8
FE4-S3	2.2705	1.058	0.54	1.0
FE1-(Cys43-S <sub><math>\gamma</math></sub> )	2.2563	1.001	0.72	2.3
FE2-(Cys46-S <sub><math>\gamma</math></sub> )	2.2824	1.034	0.61	4.6
FE3-(Cys61-S <sub><math>\gamma</math></sub> )	2.2626	0.987	0.84	0.0
FE4-(Cys75-S <sub><math>\gamma</math></sub> )	2.2709	0.944	0.73	5.0

\*Parameter  $d_1$  is the distance between the Fe atom of the pair and the BCP.

**Extended Data Table 6 | The *d*-orbital populations of iron atoms**

	$z^2$	$x^2-y^2$	$xz$	$yz$	$xy$	total
FE1	2.00 (26.4%)	1.99 (26.3%)	0.98 (12.9%)	0.65 (8.6%)	1.97 (25.9%)	7.59
FE2	1.76 (23.7%)	1.93 (26.0%)	1.98 (26.7%)	0.83 (11.2%)	0.91 (12.3%)	7.43
FE3	1.99 (28.3%)	2.00 (28.4%)	1.01 (14.4%)	1.55 (22.0%)	0.49 (7.0%)	7.04
FE4	1.99 (29.9%)	1.05 (15.8%)	0.64 (9.7%)	0.99 (14.9%)	1.99 (29.9%)	6.68



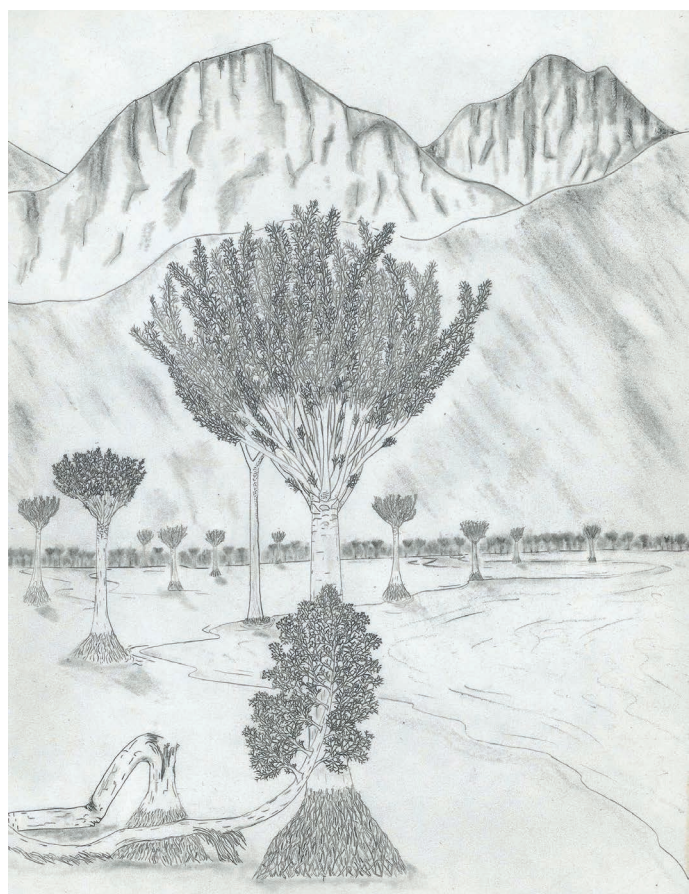
# CAREERS

**ACADEMIC ADMINISTRATION** The necessary training is there, if you ask for it **p.287**

**FACULTY WORK** Balancing demands for meetings and research [go.nature.com/20vnfif](http://go.nature.com/20vnfif)

**NATUREJOBS** For the latest career listings and advice [www.naturejobs.com](http://www.naturejobs.com)

FRANK MANNOLINI/NEW YORK STATE MUSEUM



VICTOR LESHYK



Scientific illustrator Victor Leshyk used a sketch from researchers (left) to create a conception of the Gilboa Fossil Forest for a cover of *Nature* (right).

## SCIENCE ILLUSTRATION

# Picture perfect

*Enlisting the help of an illustrator can add impact to research papers and outreach projects.*

BY JYOTI MADHUSOODANAN

**O**n canvas, a 390-million-year-old forest springs to life. Massive tree trunks jut into a sunlit clearing from a crowded forest floor. Stubby green branches battle with frilly leaf-like filaments to touch the pink-tinged sky. Palaeobotanist Chris Berry had worked for years with samples from the Gilboa Fossil Forest in New York, but had never before seen what the living forest might have looked like so many millennia ago.

Dubbed 'Lost Worlds', the digital oil painting was created by Victor Leshyk to accompany a

2012 research paper in *Nature* by Berry and his colleagues (W. E. Stein *et al.* *Nature* **483**, 78–81; 2012). It was commissioned to appear on the cover of the journal and Berry features it in his talks today, especially those for lay audiences.

It was Berry's first experience in teaming up with a scientific illustrator, and Leshyk's work exceeded his expectations. "It was very prestigious for us to have it on the cover, and the image proved very good for engagement and outreach," he says. Berry, who is based at the University of Cardiff, UK, has collaborated with artists twice since then, for press releases and museum exhibitions that involve his research,

and he is discussing a second project with Leshyk. "If you've got a story you want to get out there and you've got a really good image," he says, "it will fly a lot farther than just words."

The use of striking images to accompany manuscripts and outreach efforts is growing as more journal publishers are requiring graphical abstracts — depictions of a paper's main thrust or concept — to accompany studies. These commissioned illustrations differ from the everyday photograph, sketch or overview figure that usually accompanies research manuscripts or talks. They get to the core of concepts; they may also depict unobservable phenomena, ►



► ranging from subatomic particles to what extinct life forms might have looked like. Although working on such images with an illustrator might seem like a lot of extra toil, and paying for their services extravagant, the benefits of skilled artistic presentation can be manifold.

Visually stunning representations that result from collaborations between scientists and artists can grab millions of online views, and attract a much wider audience than a non-illustrated paper, both of which are particularly useful for researchers whose grant applications or funding proposals require them to show a public-outreach component. They are also more likely to be written about and shared digitally, helping to raise the visibility of a scientist's work, attract more students to a lab, boost career standing and improve chances of garnering funding. They can even inspire new experiments — or reveal gaps in knowledge.

Even when photographs or images already exist, hand-rendered or digital illustrations and 3D animations can clarify and enhance the technical details of a key data point or finding — exactly how proteins latch onto the surface of DNA, for example, or the shape of butterfly larvae that are usually hidden in leaf litter. Scientists who want to examine their research question or findings more fully, to 'see' their data or to provide a pictorial boost to their manuscript should consider teaming up with an illustrator. Scientific artists can also help to create artwork for a project's website, or explain hard-to-grasp concepts with short videos.

## LEARNING POINT

Most such collaborations begin when researchers are writing a paper, but it can be helpful to start even earlier (see 'Turn science into art'). Discussing with an artist how best to depict a mechanism or process — what to include and exclude, how molecules, stars or fossils should be positioned relative to one another — can help researchers to hone their hypothesis, reveal points of disagreement between authors and even identify holes in understanding.

Chemist Lauren Benz of the University of San Diego, California, found that talking with an illustrator helped her to uncover important issues that she had not considered when she started drafting her review article about the applications of membranes made from polymers and other materials. She had commissioned freelance artist Mary O'Reilly, who earned her PhD in biological chemistry from the Massachusetts Institute of Technology in Cambridge, to help illustrate how these membranes work at the molecular level. O'Reilly asked whether she should depict molecules filtering through a particular spot in the membrane, and Benz and her collaborators realized that they didn't know exactly where the filtering happened.

"It made me question some assumptions I had about the filtration mechanism, and going

back and forth with Mary helped us come up with some research questions we could ask going forward," she says. She is now planning experiments to tackle them.

Scientific illustration can encapsulate information that is not easily or often conveyed by text, line drawings or simple graphics. But it can also be used when direct imagery such as photographs are impractical or even impossible. Biologist Jessica Linton, who works with the Canadian consulting firm Natural Resource Solutions in Waterloo, was working on a recovery strategy for the endangered mottled duskywing butterfly (*Erynnis martialis*) when she realized that there were no available images of the creature's microscopic eggs and pupae, which tend to be buried in soil under leaf litter, and are extremely difficult to photograph.

Armed with scientific descriptions, she turned to illustrator Emily Damstra, whom she had met through a local butterfly enthusiasts' group. Damstra's illustrations — which are now included in the Ontario government's policy document outlining the recovery strategy — received enthusiastic appreciation from butterfly researchers and ecologists.

For those who work at the molecular level, illustrations and videos often provide the first visualization of materials or concepts that the researchers might have worked on for years — and it can be a revelation. As a graduate student, Janet Iwasa often found herself and her lab mates resorting to stick-figure drawings or waving their hands around to depict the movements of the protein they were studying: kinesin, which scuttles along skeletal filaments inside cells. "Scientific information was often lost," she says. "The first time I really understood how kinesin worked was when my principal investigator hired an animator to illustrate it." (In part because of her frustration over this, she left bench research after completing a post-doc and now works on molecular visualization in her post at the University of Utah in Salt Lake City.)

These depictions can offer surprising perspectives. "Sometimes, you need an image to tell the story effectively," says visual science communicator Kate Patterson of the Garvan Institute of Medical Research in Sydney, Australia. "They can also be question-generating, as scientists start to think about what they're seeing in a new way." When Patterson showed some researchers her animation of how DNA can be

**"Sometimes, you need an image to tell the story effectively."**

modified at the chemical level, a lively discussion ensued about how the process. Thanks to the animation, the group began to consider the physical arrangements of molecules inside the nucleus, rather than just the chemistry or enzymes involved.

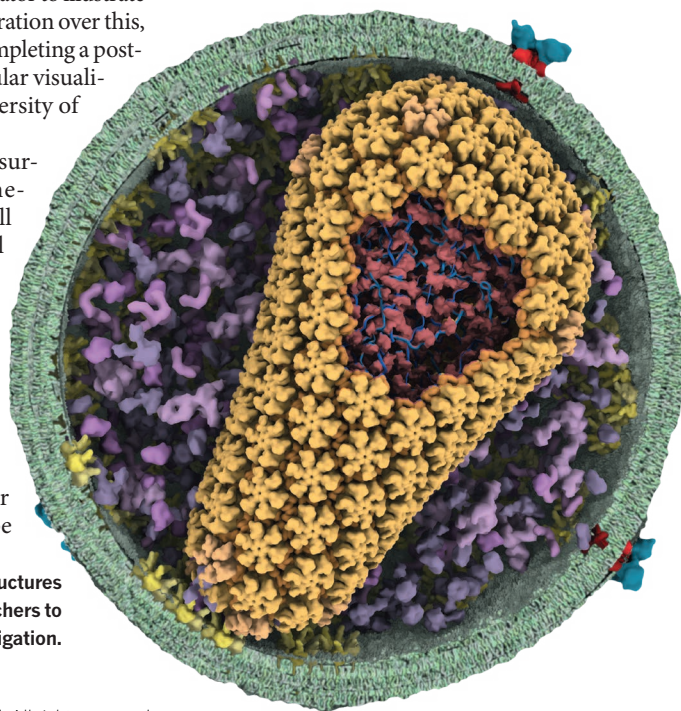
Working with illustrators can also help scientists to hone their own skills at presenting data in images. Cell biologist Matt Thomson at the University of California, San Francisco, says that collaborating with science illustrator Jessica Huppi for his study on embryonic cells taught him to prune less-relevant details for better impact, and that colour and layout can often convey information more effectively than text labels.

The paper showed that genes in growing embryonic cells can be controlled by light (C. Sokolik *et al. Cell Syst.* **1**, 117–129; 2015), and Huppi's illustrations helped him to realize that there were many ways of conveying information visually. Seeing how Huppi used effects such as colours, shapes and relative sizes has helped him to represent data effectively in subsequent work, he says. "Working with an illustrator gives you a chance to learn how to approach this type of process of thinking visually — how you convey time in a drawing, or how you can convey cause and effect."

## CONCEPTUAL APPROACH

Many researchers who have worked with illustrators say that they expect to do so again. But they note that the time needed to produce good artwork can add weeks to preparing a paper, and the expense of hiring a professional ranges from a few hundred dollars to thousands. This kind of time and money is not always defensible. Benz says that illustrations are useful for portraying general ideas or concepts, but that simple data can often be conveyed clearly in charts and graphs. Thomson

JANET IWASA



**Molecular visualizations of structures such as HIV can point researchers to new avenues of investigation.**

## GET STARTED

## Turn science into art

Here are some tips for getting the most out of the experience of creating art for science.

- Establish a working relationship with an illustrator long before you will need her or him — when you start writing a review article, for example, or when pursuing outreach projects for schools or museums.
- Seek out illustrators who have expertise in areas related to your research and look through their portfolios for artistic styles that you like. Scientists typically find artists through referrals from colleagues or through online searches for illustrators in their geographic area or field of study. The Guild of Natural Science Illustrators

in Washington DC maintains a list of contacts, and many illustrators share their own work on Twitter under #sciart.

- Clearly establish the data points that need to be in the art from the outset, so that the end product is accurate. But allow the artist to maximize the visual impact of their illustration.
- Be bold with ideas. One image isn't the definitive description of a scientific theory, so it's fine if an image includes some ambiguity about unknowns or hypotheses as long as it's done with sufficient context.
- Seek illustrators who ask questions. You should aim to find an artist who engages with your work. **J.M.**

cautions against enlisting professional help just to make a paper more decorative. Scientists who want to save money and create their own art and figures can use Microsoft Excel, molecular-visualization software and tools such as Adobe Photoshop and Illustrator, but those without artistic training may find they need to invest time in learning how to use the programs.

But they should consider more than just the money when making that choice. Hiring an illustrator saved Benz's two graduate-student co-authors from a huge time sink. "For them to not have to spend hours on learning how to draw a figure was hugely helpful," she says. "There was a direct impact on our work."

And although the right software can help a researcher to produce simple figures and visualize single molecules, that will not always result in a professional-style animation or illustration. Researchers who are not artists tend to lack the sense of design and aesthetics that are a keystones of fabulous artwork. "Where illustrators come in is in their knowledge of colour theory, using composition to guide someone's eye around a page or image in the right order," says O'Reilly. "Or drawing their eye to the centre of interest."

When Berry published a paper about a different fossil forest, his institution's press office asked him for images. With no illustrator accessible at the time, he sketched out trees by hand and sent his line drawings to a colleague who helped to add colour. The image is now widely used on websites, news stories and in research presentations, Berry says. Although his drawing was much simpler than Leshyk's, the process still took him nearly two weeks. "It was a lot of fun," he says. "But I'm not sure I could do it again. That was the first time I tried to draw a whole forest to a standard good enough to let other people look at." The experience underscored to him how much

effort — and talent — is required for illustration. Since then, he has chosen to seek professional help when he needs artwork.

Yet the value of professional scientific illustration has been tough to quantify or explain to many. Few, if any, studies have examined its impact on a manuscript, presentation or grant proposal. But many researchers vow that illustrated manuscripts get better results. "Anecdotally, people say you get more citations, or reviewers are happier with a paper, if you have good figures," Patterson says. "Or if you have a cover image, it'll get more attention. But the actual data behind that are lacking."

Still, researchers agree that whether through a simple graphic or a 3D animation, the visual communication of science is growing increasingly important. Some researchers think that professionally made figures can ease a manuscript's path through peer review. Although this is tough to verify, geneticist Deborah Kurrasch of the University of Calgary in Canada says that she has opted to work with illustrators many times before submitting a paper. And when she's acting as a reviewer, she adds, well-made figures make it easier to read and understand the data.

"Making data into art takes skill," Berry says. "If I had the resources, I would always hire an illustrator." ■

**Jyoti Madhusoodanan** is a freelance writer in San Jose, California.

## CORRECTION

The Careers Feature 'Doctor's advice' (*Nature* **533**, 429–430; 2016) incorrectly described Jelena Kovačević as a biomedical engineer. She is, in fact, an electrical engineer.

TRADE TALK  
College 'mayor'

*Elise Covic helps to design academic programmes at the College of the University of Chicago in Illinois. She explains what she does and how she gained the experience that*

*would launch her career there while she was doing her PhD in computational neuroscience.*

**What do you do as deputy dean at the college?**

When I was being recruited for this position, I was told I would be the mayor of a small, crazy town, and that I wouldn't know what problems would hit me when I woke up. I make decisions about awards for students and faculty, curricula development, development money, disciplinary actions. We have initiatives that help faculty members to help students to get the most substantive experience. I feel so lucky.

**When did you first consider this kind of career?**

By the second year of my PhD programme, I had begun to think that I didn't want a career that was research-based, but I didn't want to tell my principal investigator (PI). I was at a poster session, and my PI was proud of me, telling me I should talk to so-and-so about postdocs. Right then, I said, "We need to talk. I don't want my own lab." He said, "I don't know if I can mentor you, but let me introduce you to some friends." He directed me to the US National Science Foundation's deputy director — she invited me to call and e-mail, to come up with a plan.

**What happened next?**

I had an honest conversation with myself: what do I like to do? I love science. I love organizing. And I like to boss people around, so it was clear I could do administration. My PI said, "Why don't you run this undergraduate research programme with me?" He taught me how to administer grants and lab budgets, to deal with government agencies, to handle regulatory-compliance issues with the university. I had in-depth training. Other people could gain similar experience, if they ask the right question. ■

## INTERVIEW BY MONYA BAKER

This interview has been edited for length and clarity. See [go.nature.com/1svOHIM](http://go.nature.com/1svOHIM) for more.



# THE MEMORY WARD

*Forgive and forget?*

BY WENDY NIKEL

Liza was losing her memories. She hacked at her morning omelette and wondered what she'd lost overnight — which familiar faces had turned into strangers, which moments were gone, lost like bubbles popped in mid-air, leaving no trace of their former existence.

"Coffee?" the nurse asked. Liza couldn't recall her name. Had she known it yesterday?

"Old Liza never drinks the stuff."

Another woman with the same straight, white hair and same squarishness to her jaw shuffled into the next seat. Now here's someone Liza could never forget, leastways not while she still remembered anything at all. Every step of the way, from the cradle Gramps carved, to their wild days on the wrong side of the law, to the Memory Ward in this second-rate nursing home, Cousin Jessa had been there. "She'll have more orange juice though, if you've got some."

"Morning." Liza sighed as the nurse poured her more orange juice, orange juice she didn't even really want.

"You've got that look on your face." Jessa frowned. "You took out more memories last night?"

"Yeah, I did." Liza shrugged sheepishly. "I know it's been years, but ... I don't want to forget him."

Liza had been losing bits and pieces of her life for more than a decade now to that wretched disease. At least this way, when she extracted them into Memory Cubes, she could revisit them whenever she wanted to, even if they were gone from her mind.

Jessa scoffed. "Only a fool like you would try to save a memory by wiping it from your mind."

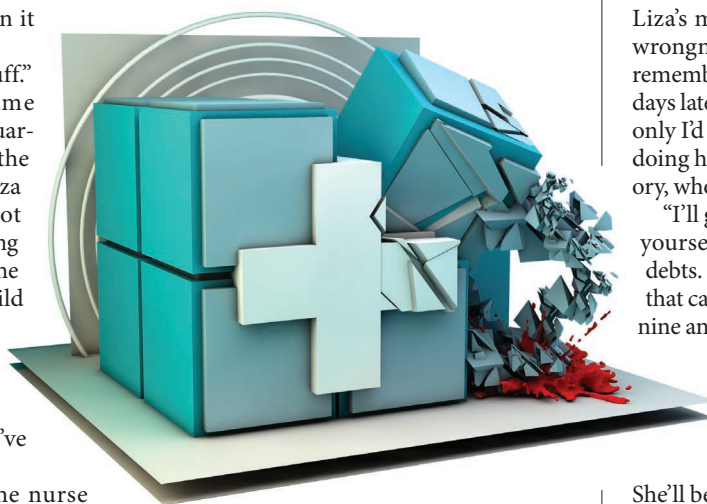
"Don't tell me you've never put anything into the Cubes."

"Course I have. Don't ask what." She shook a heap of sugar into her coffee. "That's the difference between us. You put stuff in there to remember it; I put stuff in there to forget it."

Liza must have already tucked away the memory of the first time she met him, because she couldn't recall a thing about it. She remembered later that year, though, when she told Jessa that she'd promised to go

straight, find an honest job, and take him up on his proposal. Jessa hadn't taken it too well.

"We're peas in a pod," Jessa said. "Partners in crime. And he's nothing — *nothing* — like us. You marry him and next thing you know, he'll be wanting kids and a house with a white picket fence and roast beef on the table every night. Is that really the kind of life you want?"



That was the one time in her life she stood up to Jessa; she couldn't help whom she loved.

Just thinking about him made Liza want to see him one last time, the way he was that night before their wedding, before everything went wrong ... before a sharp turn on a slick road stole away the love of her life.

Liza searched for a nurse on her way to the Cube cabinet. Somehow, although they usually hassled her all day, now that she needed one, they were nowhere to be found. She reached for the call button in her cardigan pocket, but the only thing there was the butter knife from the breakfast table. She really was losing her mind.

Never mind that, she thought, looking from the knife to the cabinet. They hadn't called her Lock-Pick Liza for nothing.

Moments later, the cabinet swung open. She scanned the Cubes, shooting a guilty glance over her shoulder. She couldn't remember the numerical code used to identify her cubes, but fortunately, they were also

engraved with the memory's date, and she'd never forget that.

She found the silver cube with the

correct date and cradled it in her hands.

Then she flicked the switch.

Suddenly she was back there again.

She sat beside him in his Jeep, but even in the dim light of passing headlights, she could tell that his face was pale and damp with tears. His hands clenched the wheel. The wipers worked furiously.

Even though it'd been rubbed out from Liza's mind, there was no mistaking the wrongness of this memory. She distinctly remembered a scene from the funeral a few days later, when she'd said over and over, "If only I'd gone with him ..." So what was she doing here now? And if it wasn't her memory, whose was it?

"I'll give you enough to make a start for yourself elsewhere. Pay off your mama's debts. Just call off the wedding." The voice that came from within her, although feminine and familiar, was not Liza's.

"I'm not interested."

"You think she really knows what she's getting into? That she'll really be content with what you have to offer her? What a joke.

She'll be crawling back to me, bored out of her mind, before a year's passed."

"I'll take that chance."

"So you won't take my offer?"

"Never."

A thin arm shot out from where Liza watched and grabbed the wheel. His cry filled the car as it swerved over the white line. The woman-who-was-not-Liza flung herself out onto the quickly passing ground.

The car crumpled into a retaining wall.

Even before the memory had fully faded, the Cube clattered to the floor. Liza hunched over, clutching her chest.

All those years ... All those lies ... How could she? How *dare* she?

"Liza?" It was the same voice as in the memory.

Liza reached for the butter knife. She'd killed before; did she still have it in her? But this wasn't just some rival crony; this was Jessa. How could she live with herself?

The glimmer of a Cube caught her eye. She gripped the knife, the answer now crystal clear. Jessa wasn't the only who could forget. ■

**Wendy Nikel** lives just west of the Wasatch Mountains in Utah. She enjoys writing speculative fiction, drinking coffee and going on roadtrips. Her website is [wendynikel.com](http://wendynikel.com).

➔ **NATURE.COM**  
Follow Futures:  
@NatureFutures  
[go.nature.com/mtoodm](http://go.nature.com/mtoodm)